IRE Transactions

ON AUTOMATIC CONTROL

PGAC-5 **JULY, 1958**

PROCEEDINGS OF THE 1957 PGAC SYMPOSIUM ON NONLINEAR CONTROL

PART I: PRACTICAL APPLICATIONS IN NONLINEAR CONTROL

Invited Papers

PART II: OBSTACLES TO PROGRESS IN NONLINEAR CONTROL

Panel Discussion

PUBLISHED BY THE

PROFESSIONAL GROUP ON AUTOMATIC CONTROL

IRE PROFESSIONAL GROUP ON AUTOMATIC CONTROL

The Professional Group on Automatic Control is an organization, within the framework of the IRE, of members with principal professional interest in Automatic Control. All members of the IRE are eligible for membership in the Group and will receive all Group publications upon payment of the prescribed fee.

Annual Fee: $2.00

IRE TRANSACTIONS®
on Automatic Control

George S. Axelby, *Editor*, Air Arm Division, Westinghouse Electric Corp., Box 746, Baltimore, Md.

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Automatic Control, 1 East 79th Street, New York 21, New York. Responsibility for the contents rests upon the authors, and not upon the IRE, the Group or its members. Individual copies available for sale to IRE-PGAC members at $0.85, to IRE members at $1.25, and to nonmembers at $2.55.

Special Symposium Issue

It may be considered that nonlinear systems are several dimensions removed from linear systems, and because of this they are more difficult to design and to understand. The nature of the input signals and changing modes of operation affect system performance in a manner that is not easily predicted with direct mathematical analyses. However, nonlinearities are always present in control systems whether intentional or not, and efforts to provide a better understanding of nonlinear system operation have been made since the origin of automatic control.

In future control systems, the use of larger automatic machines, the need for lighter missile components, the demand for greater accuracy, the vagaries of process control, and the introduction of significant nonlinearities for improved performance within bounded operating regions will emphasize the effects of nonlinear operation. The solution of the associated problems will be of increasing interest.

To assess some of the progress in this field, and to define some of the obstacles that prevent more rapid advances, the PGAC sponsored a Symposium on Nonlinear Control on August 19, 1957, in the Mark Hopkins Hotel, San Francisco, Calif. The Symposium was divided into two parts. The morning session, devoted to "Practical Applications in Nonlinear Control," consisted of three invited papers. A luncheon talk on inertial navigation preceded the afternoon session, which featured a panel discussion on "Obstacles to Progress in Nonlinear Control."

Although the one-day Symposium was comprised of only these two sessions, it attracted 114 persons. And, the audience participation was considerable in the morning session as well as during the open panel discussion in the afternoon.

Apparently, the Symposium aroused much interest, because many inquiries have been received about the papers and panel discussions. Fortunately, all of the panel discussion was recorded, and this information was transferred to typewritten copy. The papers presented in the morning session were procured, and it became possible to make the entire Symposium proceedings available to PGAC members in this special issue of TRANSACTIONS.

It should be noted that unlike other TRANSACTIONS issues, the material presented here has not been reviewed, criticized, or rewritten. This has been done in an effort to record the Symposium, as it originally took place, for those who could not attend. The readers may draw their own conclusions, formulate their own ideas, and if they wish, make their own comments, queries, and criticisms in future TRANSACTIONS.

This was the first Symposium sponsored by the PGAC with participation by the AIEE and ASME Control Committees. Although relatively small, it was well managed, the program was interesting, and the theme was significant. The committee members listed in this issue are to be congratulated for a job well done. Special acknowledgment is due to Gene Grabbe, Harold Chestnut, and their co-workers for making the panel discussion available. —*The Editor.*

The Issue in Brief

This issue of TRANSACTIONS *is arranged according to the Symposium program. The invited papers (morning session) are given in Part I, followed by the panel discussion (afternoon session) in Part II.*

PART I: PRACTICAL APPLICATIONS IN NONLINEAR CONTROL

M. V. LONG, Shell Development Company, *Chairman*

***The Design and Performance of a Model, Second-Order, Nonlinear Servomechanism,* R. E. Kuba and L. F. Kazda**—Page 43

The design and performance characteristics of a model, second-order, nonlinear servomechanism are described. The physical aspects of the various building blocks needed for nonlinear servo construction are considered. Schemes for generating nonlinear functions are given, along with methods for accomplishing the multiplication and differentiation of the control system's variables. Schematic diagrams of the model linear servo are presented, in addition to instructions for adjusting the system parameters to obtain the proper response behavior. The results of a comparison of equivalent nonlinear and linear servo operation are set forth in the form of photographs illustrating the control system's error response for various magnitudes of step inputs of position. The nonlinear system's error response is shown to be superior to the linear system's error response in that the nonlinear response exhibits no overshoot and reduces the system error to zero in approximately the time required for the linear system's error response to pass through zero for the first time. The response time of the nonlinear, model servo corresponds closely to the predicted response time.

***The Practical Realization of Final-Value Systems with Limiting Constraints,* R. C. Booton, Jr. and A. Rosenbloom**—Page 49

In certain applications, the desired function of a control system is to control some variable accurately only at one instant of time. Examples of this can be found in homing missile guidance systems or aircraft landing systems. In the latter case, it is desired to have the vertical component of velocity small in the instant of touchdown. Complicating features of such systems are the constraints imposed by physical limitations, such as the limited vertical acceleration capability of airplanes, and the noise or inaccuracy of measurement of the controlled variable. The optimum form, derived by Booton, then can motivate the design of the actual system. Some of the problems associated with the closed-loop realization of such systems are examined, such as the form, number, and utilization of measurements of the system response, how far to deviate from the ideal "bang-bang" characteristic, and methods for determining the accuracy of the practical realization.

***Combined Hysteresis and Nonlinear Gains in Complex Control Systems,* R. V. Halstenberg**—Page 51

This paper covers the origins and types of backlash, effects of backlash on positioning accuracy and limit cycles, analysis techniques using describing functions and an analog computer, elimination of limit cycles by compensating networks, and nonlinear gains and describing functions.

The combined effects of backlash and nonlinear gain (demonstrating that some unintentional low neutral gains are advantageous) are discussed for 1) backlash and nonlinear gain together in a system (they can be combined into a single nonlinearity) and 2) backlash and nonlinear gain in different parts of a system, both separated by linear elements in a single loop and in different loops of a multiloop system.

PART II: OBSTACLES TO PROGRESS IN NONLINEAR CONTROL

HAROLD CHESTNUT, General Electric Company, *Moderator*

The panel members discussed subjects concerning the session's theme, and then the session was opened to audience participation. Most of the panel discussion is preserved in this part of TRANSACTIONS. The members and their topics follow.

***Availability of Necessary Theory for the Analysis and Design of Nonlinear Systems,* O. J. M. Smith**—Page 60

The important nonlinear theory available is reviewed, and its applicability to present-day problems considered.

***Nonlinearities in Machine Tools and Missiles,* John L. Bower**—Page 62

Similarities given are between the vector space language used in machine tools and missile flight control systems. Common problems of stiction and transport lag are examined.

***Nonlinearity in Process Systems,* Ernest G. Holzmann**—Page 63

Three fundamental obstacles to progress in process control are outlined as 1) inadequacy of mathematical training, 2) lack of precise knowledge concerning dynamic and other processes to be controlled, and 3) economic factors.

***The Role of Computers in Analysis and Design of Control Systems,* George P. West**—Page 65

The use of computers in solving difficult control problems is told. It is emphasized that all problems are not easily solved, and some of the complex digital-analog computers and routines necessary to solve current control problems are described.

***Problems of Nonlinearity in Adaptive or Self-Optimizing Systems,* Charles F. Taylor**—Page 66

The adaptive or self-optimizing system is studied, along with the problems in describing and synthesizing its characteristics. It is suggested that, among other things, tradition may be a hindrance in solving advanced problems in nonlinear control.

Open Panel Discussion

A summary of the pertinent ideas discussed in the open panel discussion is included in this section.

Closing Remarks

The comments of the panel members on the nonlinear control discussions of the day conclude this special issue.

BOOK REVIEW

"*Analytical Design of Linear Feedback Controls,*" G. C. Newton, Jr., L. A. Gould, and J. F. Kaiser, *Reviewed by J. E. Gibson*—Page 74

Although this excellent review book is not related to the theme of this special issue, it is timely and interesting, and is included to prevent further publication delay.

PART I: PRACTICAL APPLICATIONS IN NONLINEAR CONTROL

INVITED PAPERS

The Design and Performance of a Model, Second-Order, Nonlinear Servomechanism

R. E. KUBA† AND L. F. KAZDA‡

Summary—This paper discusses the design and performance characteristics of a model, second-order, nonlinear servomechanism. The design procedure stems from principles laid down in the authors' previous paper, in which the nonlinear differential equation characterizing this servo was synthesized. The physical aspects of the various building blocks needed for nonlinear servo construction are considered. Schemes for generating nonlinear functions are described, along with methods for accomplishing the multiplication and differentiation of the control system's variables. Schematic diagrams of the nonlinear servo are presented, in addition to instructions for adjusting the parameters to obtain the proper response behavior. The results of a comparison of equivalent nonlinear and linear servo operation are given in the form of photographs illustrating the control system's error response for various magnitudes of step inputs of position. The nonlinear system's error response is shown to be superior to the linear system's error response in that the nonlinear response exhibits no overshoot and reduces the system error to zero in approximately the time required for the linear system's error response to pass through zero for the first time. The response time of the nonlinear, model servo corresponds closely to the predicted response time.

Nomenclature

- b_P real constants in the differential equation describing system dynamics
- E servomechanism position error
- E_1 initial value of position error
- E_{1M} maximum initial value of position error
- $f(\ )$ function of
- $f_1(\ )$ function of
- F_M coefficient of viscous friction
- G_n servo gain constants
- $h(\ )$ function of
- J_M moment of inertia of motor
- N gear ratio
- t real time
- T_M developed motor torque
- T_{C2M} response time of optimum second-order contactor servo to maximum step input of position
- T_{22} response time of second-order, class two, nonlinear servo to step input of position
- R servomechanism input variable
- C servomechanism output variable
- α_0 output acceleration
- α_M motor acceleration

† Wayne State University, Detroit, Mich.
‡ University of Michigan, Ann Arbor, Mich.

Introduction

In a previous paper,[1] the authors presented a method for the synthesis of nonlinear servomechanisms. The behavior of these synthesized, nonlinear servos, when stimulated by step inputs of position, approaches as closely as desired the behavior of optimum contactor servos of the same order. The synthesis procedure begins with the selection of a phase trajectory in a phase space, the coordinates of which are system error and derivatives of system error. Theorems were established which place sufficient conditions on the given phase trajectory so that the synthesis procedure yields a nonlinear differential equation of the form

$$\frac{d^PE}{dt^P} + h(E)\frac{d^{P-1}E}{dt^{P-1}} + b_{P-2}\frac{d^{P-2}E}{dt^{P-2}} + \cdots + b_1\dot{E} + b_0E = 0 \qquad (P \geq 2). \quad (1)$$

When the method is applied to a second-order system, the resulting synthesized nonlinear differential equation which exhibits a response time only 11 per cent greater than the response time of the optimum contractor system is

$$\ddot{E} + \frac{\pi}{2T_{22}}\sqrt{\frac{E_1}{E} - 1}\ \dot{E} + \frac{\pi^2}{2T_{22}{}^2}\ E = 0. \quad (2)$$

Realization of a Physical Model

Eq. (2) is not stated in the most convenient form for use in the construction of the physical servosystem because the time T_{22} is not a true constant but depends on the magnitude of the initial error. Moreover, to make the system response symmetrical for positive and nega-

[1] R. E. Kuba and L. F. Kazda, "A phase space method for the synthesis of nonlinear servomechanisms," *Trans. AIEE*, vol. 75, pt. 2, pp. 282–290; 1956.

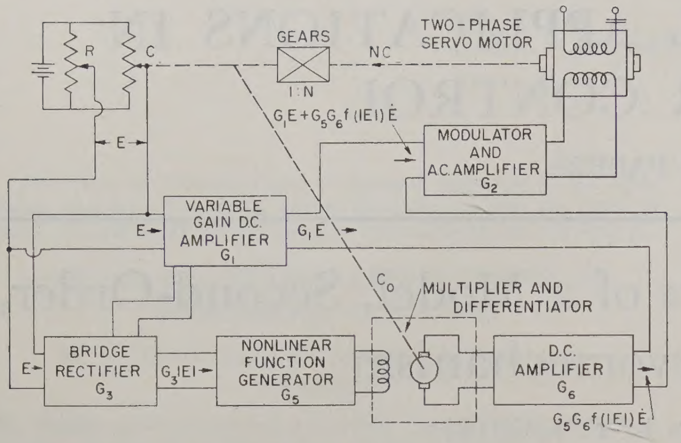

Fig. 1—Block diagram of the model, second-order, nonlinear servomechanism.

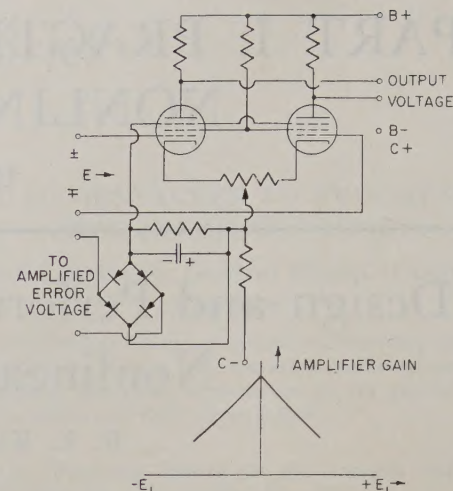

Fig. 2—Variable gain dc amplifier.

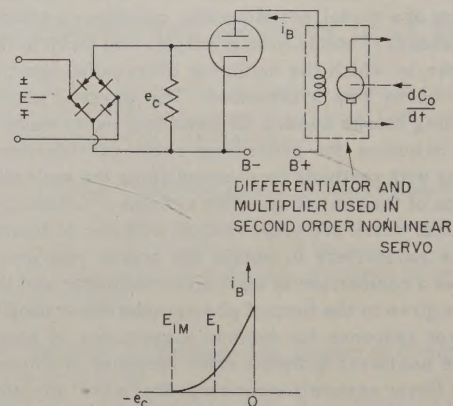

Fig. 3—Nonlinear function generator for second and third-order nonlinear servos.

tive errors, E is replaced by $|E|$ in the variable damping coefficient of (2). Therefore, the practical form of the system's differential equation reads

$$\ddot{E} + \left(\frac{\sqrt{2}}{T_{C2M}} \sqrt{\frac{E_{1M}}{E_1}} \sqrt{\frac{E_1}{|E|}} - 1\right)\dot{E} + \frac{4E_{1M}}{T^2_{C2M}E_1} E = 0. \quad (3)$$

It is evident that, since the nonlinear damping function of (3) and the coefficient $4E_{1M}/T^2_{C2M}E_1$ theoretically approach positive infinity for zero values of initial error, any practical control system with a behavior specified by (3) must terminate these quantities in positive finite values at zero initial error. Hence, the practical nonlinear servo will have an automatically built-in stable, linear mode of operation in the near vicinity of the origin of its phase space. Furthermore, the theorems of Liapounoff[2] and Bendixson[3] assure the stability of the singular point and the nonexistence of limit cycles.

Building Blocks for a Physically Realizable Nonlinear Servomechanism

Eq. (3) indicates that the important building blocks for this nonlinear servo are 1) nonlinear function generator, 2) variable gain amplifier, 3) multiplier, and 4) differentiator.

A block diagram of the model, second-order, nonlinear servo constructed by the authors is given in Fig. 1, along with the input and output quantities for each block. In this system, the error function appears as the difference of potential between the sliders of two potentiometers; one is coupled to the input shaft and the other to the output shaft. The error signal passes through two channels. In the first channel, the error signal actuates a variable gain dc amplifier. This amplifier is necessary because of the coefficient $4E_{1M}/T^2_{C2M}E_1$ in (3). A method for obtaining a dc amplifier having a variable gain determined by the magnitude of the initial error is presented in Fig. 2. Two pentodes are used as a differential dc amplifier; their suppressor grid voltages are actuated by the rectified error signal. A resistance-capacitance HOLD circuit is employed to maintain the amplifier gain at approximately the level dictated by the magnitude of the initial error. The time constant is adjusted to be nearly equal to the system response time for maximum initial error.

In the second channel, the rectified error signal is fed into a nonlinear function generator. Fig. 3 shows a scheme for producing a voltage proportional to the nonlinear damping function required in (3). Here, the rectified error signal is used to develop a negative grid voltage which changes the plate current of a triode according to its transfer characteristic. The transfer characteristic of a triode closely approximates the required nonlinear function. The output of the nonlinear function generator supplies the field current to a small dc generator coupled to the output shaft. This machine performs the operations of multiplication and differentia-

[2] A. Liapounoff, "General problems of stability of motion," *Annales de Toulouse, Paris*, vol. 9; 1907.

[3] I. Bendixson, "Curves defined by differential equations," *Acta Math., Stockholm*, vol. 24; 1901.

tion required by (3), because the armature voltage is proportional to the product of the nonlinear damping function and the derivative of the error function.

The amplified dc generator armature voltage now is combined with the output of the variable gain amplifier and fed into a modulator and ac amplifier. The ac amplifier, in turn, supplies power to the control winding of a two-phase ac servomotor. The mechanical output of the servomotor is connected by a gear train to the output shaft of the system.

System Design Equations

The equations for the servosystem illustrated in Fig. 1 are

$$E = R - C \tag{4}$$

$$G_2[G_1 E + G_6 G_5 f(|\dot{E}|)E] = J_M N^2 \ddot{C} + F_M N^2 \dot{C} \tag{5}$$

$$\frac{d^P R}{dt^P} = 0 \qquad (P \geq 1),\ (t > 0). \tag{6}$$

Thus, the nonlinear differential equation which describes the system behavior for step inputs of position is

$$\ddot{E} + \left[\frac{F_M}{J_M} + \frac{G_2 G_5 G_6}{J_M N^2} f(|E|)\right]\dot{E} + \frac{G_1 G_2}{J_M N^2} E = 0. \tag{7}$$

Now, if one compares (7) with the previously synthesized (3), the nonlinear design equations are obtained as follows:

$$G_2 G_5 G_6 f(|E|) = \left(\frac{\sqrt{2} J_M N^2}{T_{C2M}} \sqrt{\frac{E_{1M}}{E_1}} \sqrt{\frac{E_1}{|E|} - 1}\right) - N^2 F_M \tag{8}$$

$$G_1 G_2 = \frac{4 J_M N^2 E_{1M}}{T^2_{C2M} E_1}. \tag{9}$$

Engineering Design

These design specifications which were selected in order to fit the characteristics of equipment on hand are:

1) E_{1M} equal to 30 degrees of mechanical error.
2) T_{C2M} equal to 0.18 second; therefore, the response time of the nonlinear servo for a 30-degree mechanical error is 11 per cent greater or equal to 0.20 second.
3) Maximum torque on the output shaft for a 30-degree mechanical error to be 500 oz.-in.

From (9), one may calculate the maximum torque to inertia ratio required at the output shaft as

$$\alpha_0 = \frac{T_M}{N J_M} = \frac{4 E_{1M}}{T^2_{C2M}} = 64.7 \text{ radians/(second)}^2 \tag{10}$$

A gear ratio of 125:1 was selected so that the maximum required motor torque would be 4 oz.-in. The required motor torque to inertia ratio, therefore, is

$$\alpha_M = \frac{T_M}{J_M} = \frac{4 N E_{1M}}{T^2_{C2M}} = 8100 \text{ radians/(second)}^2. \tag{11}$$

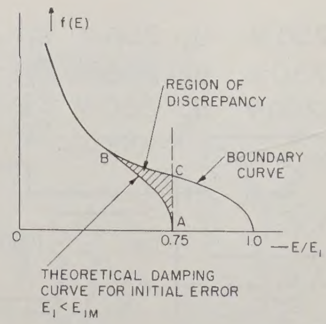

Fig. 4—Region of discrepancy between boundary curve and required damping curve for an initial error $E_1 < E_{1m}$.

A 5-watt Diehl, two-phase, 60-cycle motor was selected since its maximum output torque at full voltage would be 5 oz.-in., and its maximum torque to inertia ratio is approximately 20,000 radians/(second)^2. The moment of inertia of the motor rotor as given by the manufacturer is 0.098 oz.-in^2. The reflected moment of inertia of the 5:5:5 gear train and associated equipment was computed to be 0.034 oz.-in^2. Thus, for a maximum output torque of 4 oz.-in., the equivalent torque to inertia ratio of the motor and load is approximately 11,700 radians/(second)^2.

However, since the speed-torque curves of this motor are not vertical straight lines (*i.e.*, the motor output torque decreases as its speed increases from stall to synchronous speed), the adjusted torque to inertia ratio was calculated to be 8500 radians/(second)^2. This is in good agreement with the required value given by (11).

The voltage applied to the input and output precision potentiometers was adjusted to yield an error voltage of 0.125 volt per mechanical degree of error. The transfer constant G_2 necessary to produce 90 volts rms at the motor control winding was determined by the design specifications to be 3 volts rms output to the motor control winding per mechanical degree of error.

The variable gain constant G_1 was selected to give the same output for 3 degrees of error as for 30. Therefore, G_1 must change in the ratio of ten to one.

The term $G_5 f(0)$ in (8) was determined experimentally to be 0.0036 volt-second/radian by running a magnetization curve on the dc generator.

In (8), the term

$$\sqrt{\frac{E_{1M}}{E_1}} \sqrt{\frac{E_1}{|E|} - 1}$$

was approximated by a function $f_1(E)$. The value of $f_1(0)$ then was chosen to be 4.00 because this value had produced good results in the analog computer study reported.[1] Although a different nonlinear damping function is theoretically required for each value of initial error, all of these damping function curves climb up to the boundary curve very quickly as the error proceeds toward zero. Therefore, in this scheme, only the boundary curve is used to approximate the nonlinear damping function for all values of initial error. Fig. 4

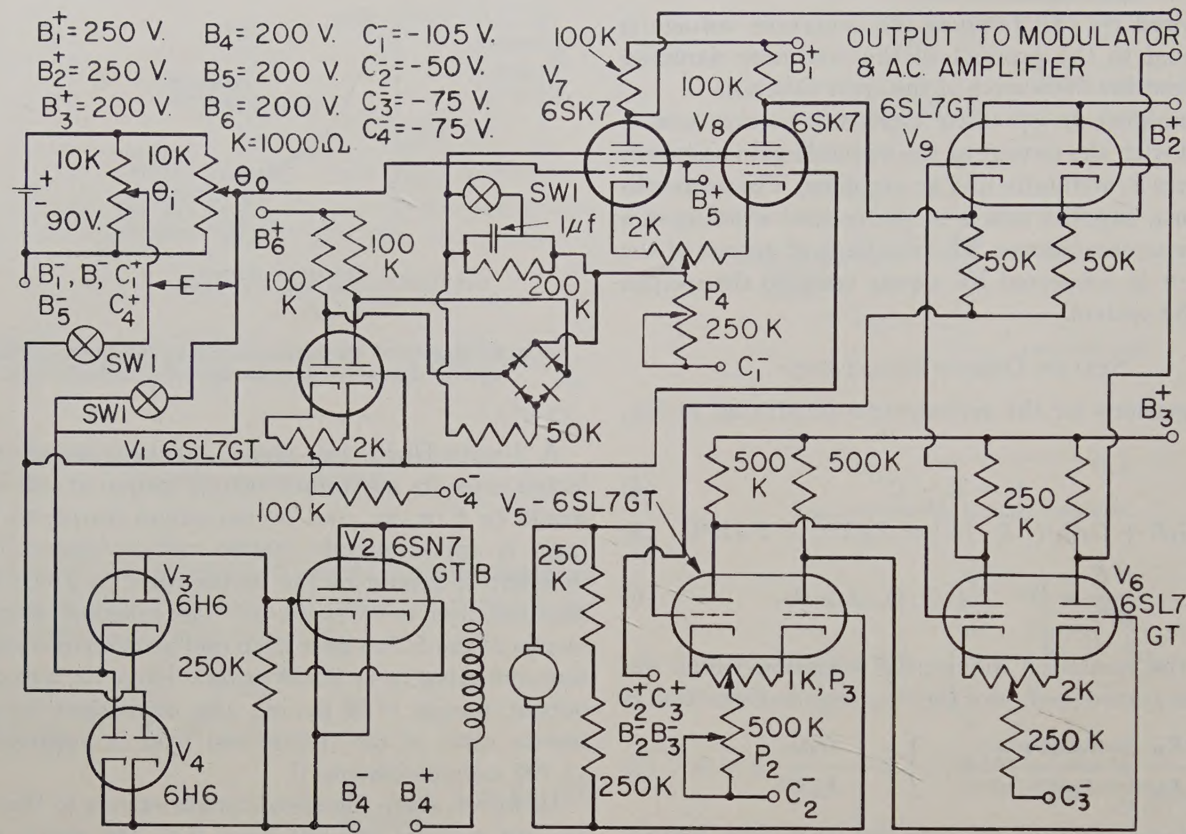

Fig. 5—Schematic diagram for the second-order nonlinear servomechanism: Part I.

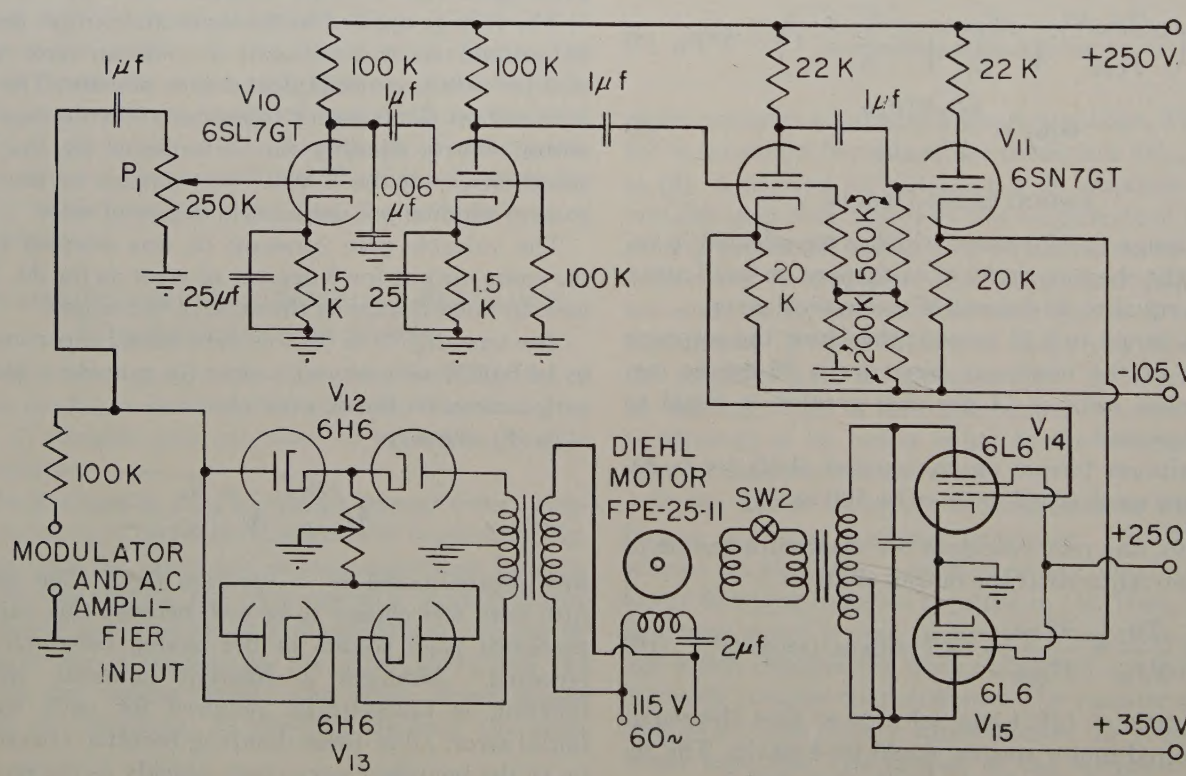

Fig. 6—Schematic diagram for the second-order nonlinear servomechanism: Part II.

presents the region of discrepancy between the boundary curve and the required damping curve for an initial error $E_1 < E_{1M}$. This approximation is justified because in the region of greatest discrepancy, the magnitude of the error derivative which multiplies the nonlinear function is small and, hence, the contribution of this product term to the operation of the system is small in the region of disagreement.

Eq. (8) now yields the value for the dc amplifier gain as

$$G_6 = \frac{\sqrt{2} J_M N^2 f_1(0)}{T_{C2M} G_2 G_5 f(0)} = 1520. \tag{12}$$

Eq. (12) neglects the effect of output viscous damping which would reduce the gain coefficient somewhat.

Figs. 5 and 6, p. 46, are schematic diagrams for the second-order nonlinear servomechanism built and tested by the authors. The error signal appears as the voltage difference between the sliders of two linear potentiometers, one coupled to the input shaft and the other to the output shaft. The error signal passes through two channels. In the first channel, the rectified error signal is applied to the grid of a 6SN7GTB triode which performs the operation of generating the nonlinear damping function. Fig. 7 shows a comparison of the theoretically required nonlinear function and the actual nonlinear function generated by the triode. The output of the triode is fed into the shunt field winding of a small dc generator coupled to the servo output shaft. The time constant of the field circuit was made negligibly small. The dc generator armature voltage, which is proportional to the product of the nonlinear damping function and the system error rate, is amplified and added to the error signal which has passed through a variable gain amplifier in the second channel. The combined dc signal actuates a balanced modulator where it is converted into an alternating voltage. The amplified ac signal is applied to the control winding of the two-phase servomotor. This motor drives the output shaft through a gear train having a 125:1 ratio.

Potentiometers P_1 and P_2 are employed to adjust the gains of the ac and dc amplifiers, respectively. In order to obtain symmetrical behavior of the nonlinear function for positive and negative errors, the balancing potentiometer P_3 is adjusted so that the plate current of V_2 reads a maximum for zero system error. The level at which the nonlinear damping function generator operates is adjusted by changing the plate voltage applied to V_2.

Linear Operation

The particular choice of building blocks used in the construction of this servo was influenced greatly by the authors' desire to be able to easily convert the nonlinear system into an equivalent linear system for a basis of comparison. This is accomplished by the operation of switch SW 1 in Fig. 5, which converts the variable gain amplifier into a constant gain amplifier and at the same time produces a constant damping coefficient. Equivalent linear operation was obtained by adjusting the amplifier gains and the damping coefficient such that at an error of 30 mechanical degrees the maximum output torque of the linear system was equal to the maximum output torque of the nonlinear system.

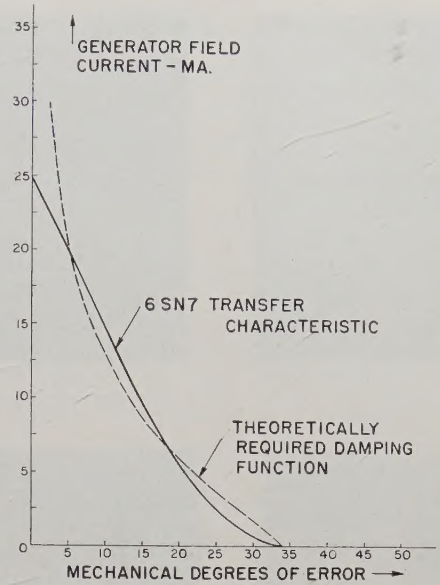

Fig. 7—Comparison of damping functions.

Experimental Results

A comparison of the error responses of the model servosystem was made by the authors for various magnitudes of step input of position under both linear and nonlinear operation. The error voltage of the system was recorded by photographing its trace on an oscilloscope when the system was subjected to a step input of position. The switch SW 2 in Fig. 6 was opened and the input shaft rotated a certain number of degrees. Next, SW 2 was closed and at the same instant the sweep of the oscilloscope energized so that the error signal could be photographed. These photographs are presented in Fig. 8 and Fig. 9. Experimental results clearly indicate that the nonlinear system possesses the following characteristics:

1) The error response is symmetrical for positive and negative errors.
2) The error response exhibits negligible overshoot.
3) The steady-state error is approximately zero.
4) The error response is superior to the equivalent linear system's error response for all magnitudes of initial error less than or equal to 30 degrees.
5) The system is stable for all magnitudes of initial error greater than 30 degrees; its error response is

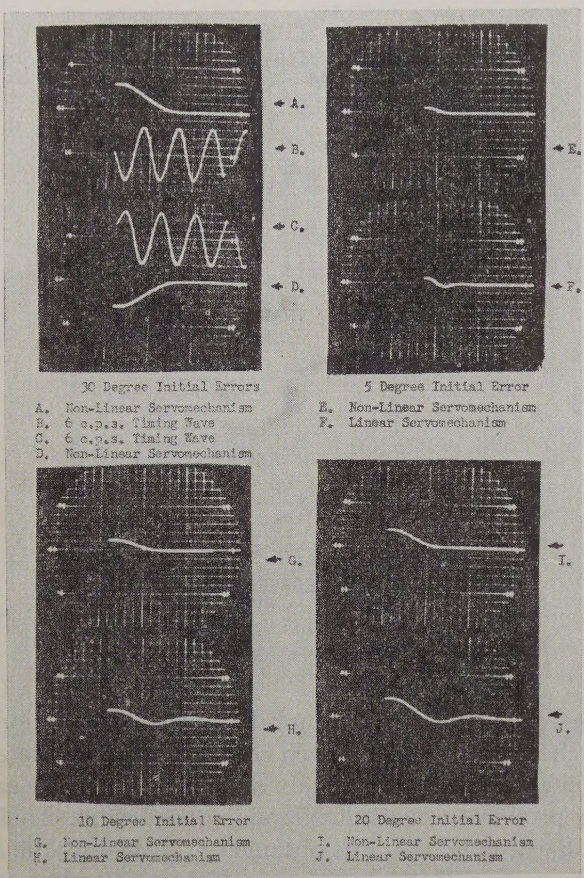

Fig. 8—Error response of the equivalent nonlinear and linear model servomechanisms: Part I.

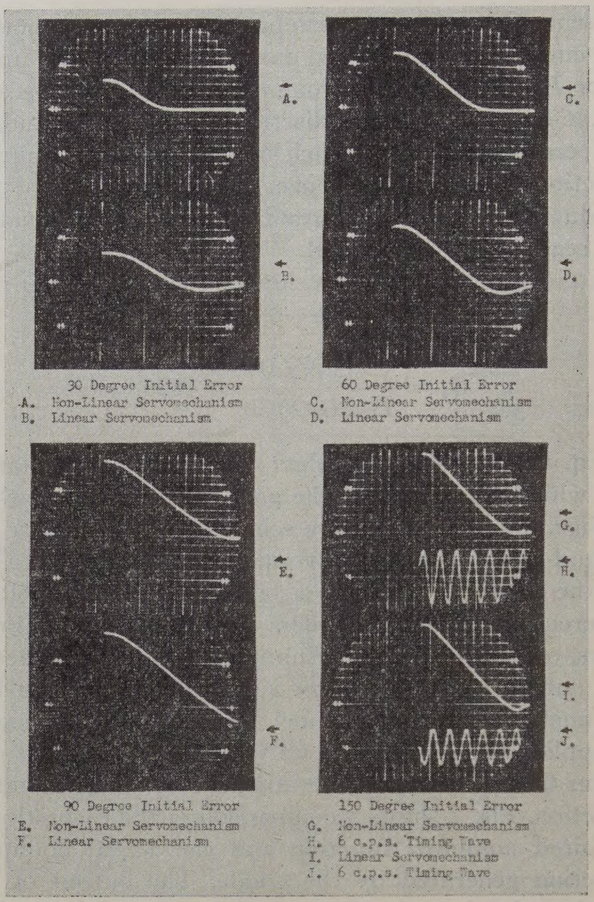

Fig. 9—Error response of the equivalent nonlinear and linear servomechanisms: Part II.

superior to the linear system's error response for these values of initial error.

6) The measured response time for a 30-degree initial error (0.213 second) is in good agreement with the theoretically predicted response time (0.200 second). The difference probably is due chiefly to the approximation used for the nonlinear damping function.
7) The response time becomes smaller as the magnitude of the initial error decreases, varying approximately as the square root of the initial error magnitude.

Other Types of Inputs

In this study, the authors have concentrated their efforts on producing a physical, nonlinear servosystem whose behavior would be in close agreement with the results as predicted from synthesis procedure. Since the synthesis procedure was based on a step input of position stimulus, the main endeavor was to establish the validity of this synthesis method to the highest degree possible. Therefore, the performance of the nonlinear servo for inputs other than step inputs was viewed by the authors in a qualitative way only. However, it was observed that 1) the nonlinear servo performed adequately well and superior to the linear system for inputs of the ramp type and for recurrent and overlapping inputs of the step and ramp type, and 2) no input was discovered which would render the nonlinear servo unstable.

Extension to Higher Order Systems

It is apparent now that the procedure for translating the synthesized nonlinear differential equation into a physical system can be easily extended to higher order systems. The important building blocks are variable gain amplifiers, multipliers, function generators, and differentiators. The design equations are given in the authors' previous paper.[1] The ingenuity and resourcefulness of servo designers is all that is needed to produce the physical, higher order, nonlinear systems with nearly optimum dynamic characteristics.

Conclusions

The performance characteristics of the model nonlinear servo discussed substantiate the authors' thesis that the synthesis approach to nonlinear problems is a most fruitful one. There is excellent agreement between the predicted performance as obtained from a straightforward synthesis procedure and the physical performance of the model servo. The nonlinear servo with its excellent dynamic characteristics can be subjected now to the ingenuity of servo designers.

The Practical Realization of Final-Value Systems with Limiting Constraints

R. C. BOOTON, JR.† AND A. ROSENBLOOM†

Summary—**In certain applications, the desired function of a control system is to control some variable accurately only at one instant of time. Examples of this can be found in homing missile guidance systems or aircraft landing systems. In the latter case, it is desired to have the vertical component of velocity be small at the instant of touchdown. Complicating features of such systems are the constraints imposed by physical limitations, such as the limited vertical acceleration capability of airplanes, and the noise or inaccuracy of measurements of the controlled variable. The optimum configuration for such systems has been derived, and this form can be used as the basis for the design of actual systems. The theoretical derivation of the optimum system idealized the actual situation, and a number of problems must be considered when the physical system is designed. Such problems as the closed-loop realization, the number of measurements of the system response, and the manner in which these measurements are utilized are discussed in this paper.**

THE final-value system is a control system whose function is to control accurately the value of a variable at one instant of time. If the physical system imposes constraints on maximum rates of change of controlled variables, the variable of interest also must be controlled in some fashion at earlier times, in order that the error in the final value be small. Considered here is a hypothetical application to the yaw control of an aircraft landing system as illustrated in Fig. 1. The radar system is assumed to measure the lateral distance $(M-R)$ of the landing aircraft from the runway. The constant M is the initial lateral distance of the aircraft from the runway when the system commences operation and R is the response (lateral displacement of the aircraft). The control problem is the generation of suitable corrective signals to the aircraft autopilot on the basis of the radar measurements so that the lateral error is small at the instant of touchdown. The lateral maneuverability of the aircraft is considered to be finite.

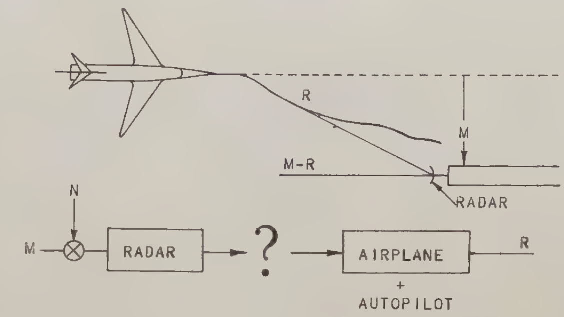

Fig. 1—Yaw control of aircraft landing system.

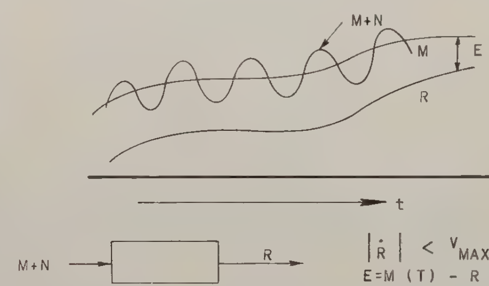

Fig. 2—Mathematical formulation of control problem.

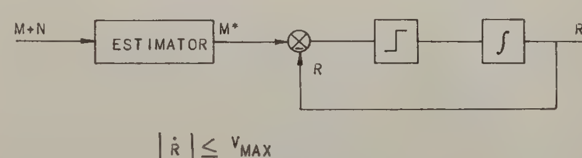

Fig. 3—Optimum form of control system.

Fig. 2 shows a mathematical formulation of this problem for the situation in which the lateral velocity $\dot{R}$ is constrained to be less than a maximum value $V_{\max}$. The radar measurement is corrupted by radar errors and is written as $M+N-R$. The optimum form of such systems has been established,[1] under the conditions that the message M and the noise N are jointly Gaussian random processes and the error criterion is an even and nondecreasing function of the absolute value of the error. This optimum form is presented in Fig. 3. The system is composed of an optimum estimator (filter) of M based on $M+N$ without regard to constraints followed by a "bang-bang" first-order control system. Thus, the second portion of the system minimizes the error between the system response R and M^* (the best estimate of M). The total system error is the sum of the estimation error $M-M^*$ and the servo error M^*-R.

The basic form of the system with a velocity constraint can be modified for other constraints by making use of an extrapolation process. An extrapolated response is so defined that it is itself velocity constrained and equals the actual response at the final time. Fig. 4 is the optimum system for an acceleration constraint. The extrapolated response R_E for this case is determined from

† Space Technology Labs., Div. of The Ramo-Wooldridge Corp., Los Angeles, Calif.

[1] R. C. Booton, Jr., "Optimum design of final-value control systems," *Proc. Symp. Non-Linear Circuit Analysis*, pp. 233–241; April, 1956.

—, "Final-value systems with Gaussian inputs," IRE TRANS. ON INFORMATION THEORY, vol. IT-2, pp. 173–175; September, 1956.

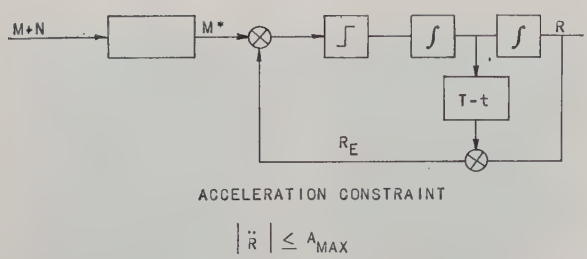

Fig. 4—Optimum system for an acceleration constraint.

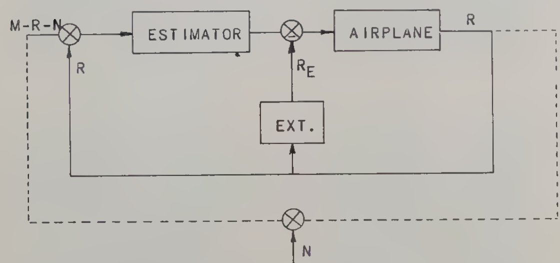

Fig. 5—Closed-loop realization of final-value system.

$$R_E = R + (T - t)\dot{R}$$

whereby

$$\dot{R}_E = (T - t)\ddot{R}.$$

Thus, if R is acceleration constrained, R_E is velocity constrained. Because $R_E(T) = R(T)$, the system which minimizes the final extrapolated error also minimizes the actual error.

The statement of the problem above has assumed that the system has available a noisy measurement of the message M, so that its input is $M+N$. In many practical situations, the measuring devices inherently measure an error signal $M-R$, so that the system input is $M-R+N$. Mathematically, this situation is made equivalent to the one originally considered by addition of $M-R+N$ and the response R, as indicated in Fig. 5. A direct method of obtaining the system response R is to measure it separately. Fig. 6 illustrates a system in which the response is measured using a double integrating accelerometer. The system is mechanized here with a one-way communication link from the ground to the aircraft.

An accurate measurement of the system response is not always possible, and it may not be desirable even

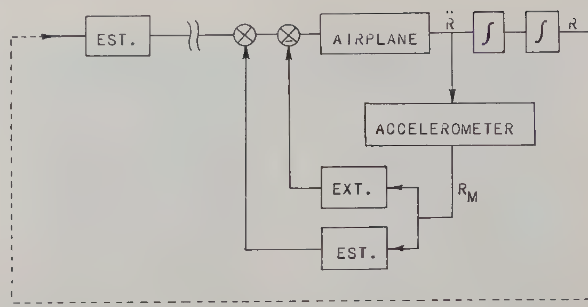

Fig. 6—Response measurement using double integrating accelerometer.

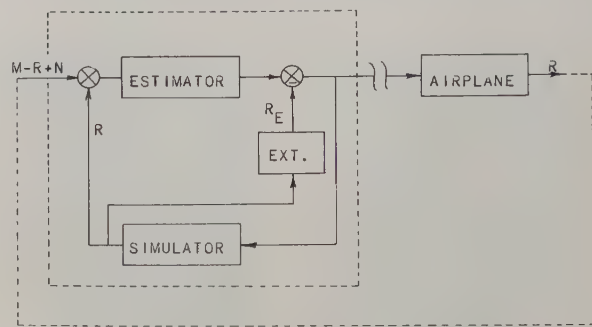

Fig. 7—Final-value control with simulator.

when it is. If the characteristics of the controlled element are known accurately enough, a signal representing the response may be found without measurement. This process, here called simulation, consists of nothing more than a calculation of the response of the assumed controlled element to the same forcing function that is applied to the actual controlled element. The final-value system resulting from this approach is seen in Fig. 7. Here, the simulation and extrapolation computations are performed by a ground-based computer, and steering commands are transmitted to the aircraft through a communication link.

The final-value concept has proved of great value in the design of certain classes of control systems. The ideal configuration usually must be modified to effect a reasonable mechanization. Even then, the increase in accuracy over conventional techniques can be considerable. Although this paper has mentioned a few of the problems encountered in the design of actual systems, much work remains to be done, in particular, error analyses of ideal systems and practical realizations.

Combined Hysteresis and Nonlinear Gains in Complex Control Systems

R. V. HALSTENBERG†

Summary—The types and sources of hysteresis commonly found in mechanical systems, and some of the effects thereof, are discussed.

The describing function technique as specifically applied to studying hysteresis-caused limit cycles is outlined.

The use of an analog computer for studying hysteresis in servosystems is reviewed.

Techniques are presented for determining active and passive compensation to eliminate or reduce hysteresis limit cycles.

Finally, a method is given for analyzing hysteresis plus a single-valued nonlinearity when they occur together or apart in a single or multiloop control system. By this method it is possible to synthesize a nonlinear gain characteristic which eliminates or reduces a hysteresis limit cycle. The validity of the method is demonstrated by comparison with analog computer results.

Introduction

In the design of many closed-loop systems, it is assumed that the system is linear within the operating range. If the saturation levels that invariably exist in practical hardware are made large enough, the assumption of linearity is usually quite good at moderately large amplitude operation. However, experience has shown that at low operating levels, small amplitude, unintentional nonlinearities often render attempts at linear analysis useless, particularly where mechanical components are involved.

Hysteresis

One very important low amplitude nonlinearity that is both quite bothersome and quite prevalent is known as "hysteresis." Hysteresis often is defined as "the difference between the increasing and the decreasing output." Fig. 1 illustrates the particular hysteresis characteristic, sometimes called "backlash," which is common in mechanical systems and which is dealt with at length in this paper.

The characteristic of Fig. 1 may be generated, of course, by play in gears, bearings, and mechanical connections. Perhaps even more important, since it is not always as obvious, is the fact that friction can, and often does, result in hysteresis. Whenever force is the input to a section of a system containing friction, a hysteresis loop is formed whose width is equal to twice the friction. Another friction case exists when remote positioning is attempted by means of "zero force linkages," as in some aircraft systems. A plot of input vs output for a push rod with a finite spring constant and no other axial force than output friction results in a hysteresis loop whose width is twice the friction divided by the spring constant.

† Convair, San Diego, Calif.

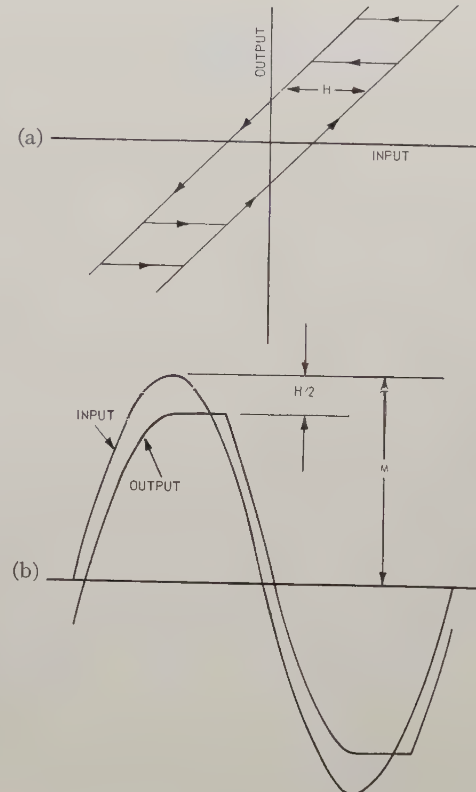

Fig. 1—Hysteresis characteristics.

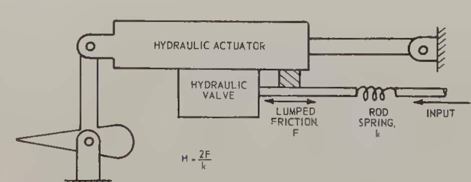

Fig. 2—Hydraulic surface positioning system.

A special case of "zero force linkage" hysteresis in aircraft systems occurs when the push rod is the input to a hydraulic-powered control surface positioning loop as shown in Fig. 2. For the input shown, the input rod spring must compress to overcome valve friction, but when the mechanical feedback attempts to close the valve, the friction force is reversed and the input rod spring is expanded. Detailed consideration of this situation reveals that the resulting hysteresis is effectively in the valve itself and not in the input. The existence and effective location of this hysteresis has been confirmed by exhaustive tests on the early F-102A elevon system, where remote input minus actuator position (actual error) was plotted vs actual valve position (effective error). The resultant curve was a well-defined hysteresis

loop. Assuming that valve force is a constant friction is not rigorously correct, since flow forces and acceleration forces also are present. But for given valve amplitudes and frequencies of importance, the net effect of all forces apparently can be represented satisfactorily by friction.

Effects of Hysteresis

The worst result of hysteresis in a system is a sustained oscillation of limited amplitude known as a limit cycle. Fig. 1(b) illustrates the basic effects of hysteresis on a closed-loop system: phase lag and gain reduction. As amplitude decreases, the phase lag increases and the gain decreases. If the system containing the hysteresis is such that one or more of these phase-gain combinations causes instability, a limit cycle may be maintained at the greatest amplitude which just gives marginal stability.

It should be kept in mind that hysteresis also affects steady-state accuracy; there is no output for an input equal to one half the width of the hysteresis. Methods are presented by which limit cycles can be eliminated, but none of them improves the steady-state accuracy.

The amplitude of limit cycles may vary widely in different systems. It is possible for the amplitude of oscillation to be such that the input to the hysteresis is only slightly larger than the hysteresis itself. In this case, the oscillation may be within the desired steady-state accuracy and possibly tolerable. On the other extreme the limit cycle may be many times the size of the hysteresis, and peak-to-peak amplitudes of five or ten times the hysteresis are entirely possible in systems that are satisfactorily stable at the larger amplitudes.

Describing Function Analysis

The analytical method of describing functions may be utilized to determine the low-amplitude stability characteristics of most systems containing a single hysteresis loop. When applying the hysteresis describing function, the following assumptions are made.

1) When the system is driven sinusoidally the higher harmonics generated by the hysteresis are effectively filtered out, and only the fundamental frequency component travels the closed loop to reenter the hysteresis.
2) Except for the one hysteresis, the system is linear.
3) The characteristics of the hysteresis are functions of amplitude only and are not frequency sensitive.

As is seen later, all of these conditions need not always be met in order for the analysis to yield practical results, but the method as outlined here cannot be rigorously justified unless the assumptions are met.

Having satisfied the assumptions, the steady-state sinusoidal transfer function of the hysteresis may be evaluated for any input amplitude by dividing the Fourier fundamental of the output by the input. This equivalent transfer function, hereafter called N_H, is given in Table I. In preparing a block diagram of the system, N_H is located at the appropriate position and the equations for a frequency response analysis written in the usual manner. From the block diagram of Fig. 3, the open-loop transfer function is

TABLE I

Describing Function for Hysteresis

H/M	N Gain	N Gain-db	N Phase °	$-1/N$ $\frac{1}{\text{Gain}}$	$-1/N$ $\frac{1}{\text{Gain}}$ db	$-1/N$ Phase °
0	1.000	0	0	1.000	0	−180
0.1	0.983	−0.15	−3	1.017	0.15	−177
0.2	0.957	−0.38	−6.5	1.045	0.38	−173.5
0.3	0.924	−0.68	−9.8	1.082	0.68	−170.2
0.4	0.885	−1.06	−13	1.130	1.06	−167
0.5	0.840	−1.52	−16.4	1.191	1.52	−163.6
0.6	0.793	−2.02	−19.7	1.262	2.02	−160.3
0.7	0.744	−2.57	−23.0	1.344	2.57	−157
0.8	0.693	−3.19	−26.2	1.444	3.19	−153.8
0.9	0.640	−3.88	−29.5	1.563	3.88	−150.5
1.0	0.590	−4.58	−33	1.695	4.58	−147
1.1	0.535	−5.44	−36	1.870	5.44	−144
1.2	0.480	−6.38	−39.4	2.08	6.38	−140.6
1.3	0.425	−7.46	−43	2.36	7.46	−137
1.4	0.366	−8.72	−46.7	2.73	8.72	−133.3
1.5	0.310	−10.18	−50.7	3.23	10.18	−129.3
1.6	0.250	−12.04	−55	4.00	12.04	−125
1.7	0.190	−14.42	−59.5	5.26	14.42	−120.5
1.8	0.125	−18.06	−65.3	8.00	18.06	−114.7
1.9	0.061	−24.30	−73	16.4	24.30	−107
2.0	0	−∞	−90	∞	∞	−90

$$KG(j\omega)N_H(M),$$

and the critical equation for the denominator is

$$KG(j\omega)N_H(M) + 1 = 0.$$

To solve this equation, the amplitude dependent portion is separated from the frequency dependent portion.

$$\left[KG(\omega j) = -\frac{1}{N_H(M)}\right]$$

and both sides of the equation plotted on an amplitude ratio-phase plot. The plot takes the same significances as a Nyquist plot, in that when a portion of the describing function is enclosed, a limit cycle is indicated at the intersection having the largest hysteresis input amplitude, provided that the odd higher harmonics of the frequency of intersection are suitably attenuated by the frequency sensitive portion of the loop. Appendix I contains an example of describing function application.

Since the higher harmonics are never entirely removed by the filtering action of the system, the validity of the describing function method is sometimes questioned. However, the harmonic output of the hysteresis is low to begin with and slight distortion of the input does not change the equivalent transfer appreciably except at inputs approximately the size of the hysteresis. If properly applied, this method will yield excellent results when the nonlinearity is hysteresis. The author has seen many cases where the frequency analysis of hysteresis alone has been borne out by computer simulations and in no case have the results been misleading.

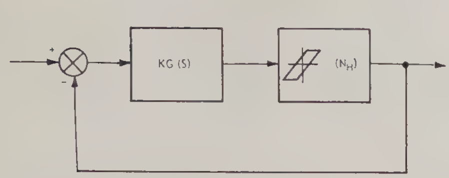

Fig. 3—Block diagram for system with hysteresis.

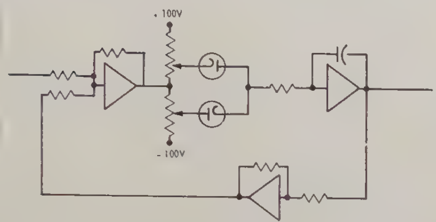

Fig. 4—Computer circuit for hysteresis simulation.

Caution should be exercised in attempting to interpret the results of the hysteresis describing function too accurately. If the two curves are nearly parallel and the intersection cannot be determined accurately, nothing is to be gained by larger, more accurate plots since minor variations in the linear system may shift the intersection within the indefinite region. This is especially true at the very small amplitudes where the inverse describing function $[-(1/N_H)]$ has a large change in gain for a small change in phase shift. Since the linear portion of the system is seldom known to an extremely high degree of accuracy, the prediction may be thought of as being limited in accuracy by the physical situation rather than the method of analysis.

Analysis by Analog Computer

When analyzing a system with hysteresis, an analog computer simulation is a valuable aid. Determining the effects of many possible changes in the linear system is much faster on the computer, and information about the transient response that is not available by describing functions can be obtained readily. If there are a number of nonlinearities in the system, there is no other general method than the computer available. The foregoing is by no means intended to belittle the describing function. It is of great value to have both the computer and analytical studies for cross checks and the analytical method has the great advantage of indicating methods of compensation.

To simulate hysteresis on the computer, a dead spot is placed before the integrator in a three-amplifier time lag circuit, as seen in Fig. 4. Since this circuit will have a linear time lag, the loop gain must be at least ten times the highest frequency (radians per second) studied and 100 times is to be preferred. This high gain can be used only if the dead spot is indeed a true dead spot, since even an extremely low gain portion destroys the sharp corners and introduces extra low frequency phase shift. A very good check is to graphically calculate the desired output from the input, as is done in Fig. 1(b). The circuit in Fig. 4 has given the author excellent results.

Elimination of Hysteresis Limit Cycles by Linear Methods

Just as a linear system can be stabilized or improved by the addition of appropriate compensation, it is possible to improve the low-amplitude nonlinear response by proper shaping of the linear locus. As might be expected, however, the nonlinear problem is not as simple as the linear one: there is a locus to be avoided instead of a single point.

The most effective compensation is the lead network $(1+\tau s)$ which is positioned frequency-wise to reduce the phase lag in the limit cycle region. Much the same effect is achieved by a lead-lag network

$$\left(\frac{1+\alpha\tau s}{1+\tau s},\ \alpha > 1\right)$$

which is usually easier to mechanize and is less subject to noise problems. A lag-lead network

$$\left(\frac{1+\alpha\tau s}{1+\tau s},\ \alpha < 1\right)$$

may be effective in some nonintegrating loops where additional phase lag can be tolerated at low frequencies but where attenuation is needed at higher frequencies. In an integrating loop the lag-lead network may be used to reduce the amplitude and frequency of the limit cycle, but only in rare cases is elimination of the oscillations possible.

Examples of the above types of compensation are given in Appendix II.

As intimated previously, another way to eliminate (or reduce) a hysteresis limit cycle is adjustment of the loop gain. The log of amplitude vs phase plot is particularly convenient for this because the frequency locus can be moved down until the describing function is no longer enclosed. However, gain usually is selected in view of some system response requirement and it generally is not practical to compromise performance at all amplitudes in order to improve the low amplitude situation.

Nonlinear Gains and Describing Functions

It is worthwhile to pause at this point and cover briefly the describing function as applied to nonlinear gains. The assumptions given for the hysteresis describing function still apply, and the gain for any sinusoidal input is expressed as the Fourier fundamental of the output divided by the assumed sine wave input. Since a nonlinear gain by itself does not produce any phase shift, to predict a limit cycle it is necessary only to determine if any of the amplitude-sensitive gains are equal to the linear gain margin.

If a low-amplitude nonlinear gain is intended, the characteristic shown in Fig. 5 is usually reasonable to obtain by means of diodes, or in hydraulic valves by shaping the valve orifices. The describing function for Fig. 5 is

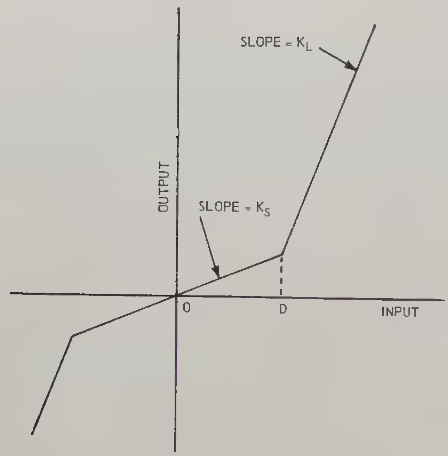

Fig. 5—Nonlinear gain.

$$K_{DF} = K_L\left[1 - \left(1 - \frac{K_S}{K_L}\right)\left(\frac{2\lambda + \sin 2\lambda}{\pi}\right)\right], \quad A > D$$

where $\lambda = \arcsin D/A$ radians and A is the zero to peak magnitude of the sine wave input. For $A<D$, $K_{DF}=K_S$. A quick approximation to the describing function for this or any nonlinear gain may be obtained by dividing the peak value resulting from a sinusoidal input by that input. The approximation is reasonably good except in the immediate region of any sudden gain change.

Hysteresis and Nonlinear Gain in Series

If hysteresis and a nonlinear gain element appear together in a system, as in Fig. 6, one approach to a describing function analysis would be to determine a describing function for the combination. This is found to be quite laborious and the result is good for only the one particular gain characteristic.

An easier and more flexible method is outlined below.

1) Prepare an amplitude-phase plot for the linear portion of the system and apply the hysteresis describing function in the usual manner.
2) From the combined plots, information can be obtained about the frequency, gain required to cause a limit cycle at that frequency, and amplitude of the limit cycle which could occur.
3) The amplitude of the possible limit cycle at the input to the nonlinear gain element is determined by multiplying the oscillating input to the hysteresis as found in 2) by the hysteresis "gain" at that particular amplitude.
4) The gain to just cause a limit cycle is plotted vs the gain element input amplitude. Various gain characteristics can then be placed on the same plot and their effects on low-amplitude stability studied.

It should be noted that the input amplitude to the gain element as obtained in step 3) is the Fourier fundamental of the input and not the actual input. Since the fundamental is used as the input, two ques-

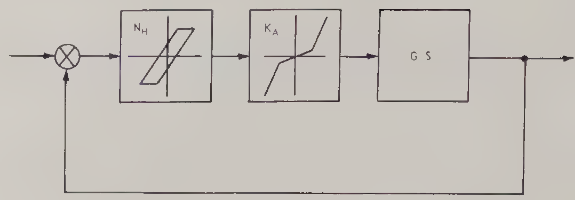

Fig. 6—System with hysteresis and nonlinear gain in series.

tions must be answered: How does the actual effective input differ from the fundamental, and what added effects will the nonlinear gain have? At all but the lowest amplitudes the harmonic content of the hysteresis output is low and the harmonics are of such phase relationship so as to make the results conservative. For example, at $H/M=1.0$, the third harmonic is 15 db below the fundamental and the peak value of hysteresis output is $0.5H$ instead of $0.59H$ as given by the first harmonic. If the nonlinear gain is such that small amplitude changes result in only small gain changes, no appreciable phase error results since hysteresis phase shift is determined principally by the time of zero amplitude. In fact, figuring gain by peak amplitudes and phase shift from time of zero amplitude gives a good approximation to the describing function determined by Fourier analysis and given in Table I. An example of hysteresis and nonlinear gain in series is presented in Appendix III, and excellent agreement with computer results is shown.

A careful inspection of the curves and computer results for Appendix III will reveal one very important fact. If a system has a limit cycle problem, a low-amplitude gain loss such as commonly found in hydraulic servo valves may be actually helping rather than harming low-amplitude stability. The effect of the gain at low amplitudes can be determined only by an intelligent analysis. Caution is directed, however, to the fact that the reduction in frequency which usually accompanies a reduction in gain may result in an increased gain in the frequency sensitive elements so the net effect must be studied at the point in the system where control is desired.

Hysteresis and Nonlinear Gain in Different Parts of a System

If the hysteresis and nonlinear gain do not appear in direct series, the system can be analyzed by the same general approach but the specific steps will depend on the particular system. A general outline which includes both separated and combined hysteresis and gain follows.

1) From the block diagram write the frequency response equations including the describing functions.
2) Manipulate the equations to isolate N_H.
3) Determine the value of gain to just cause a limit cycle and the limit cycle amplitude at each frequency.

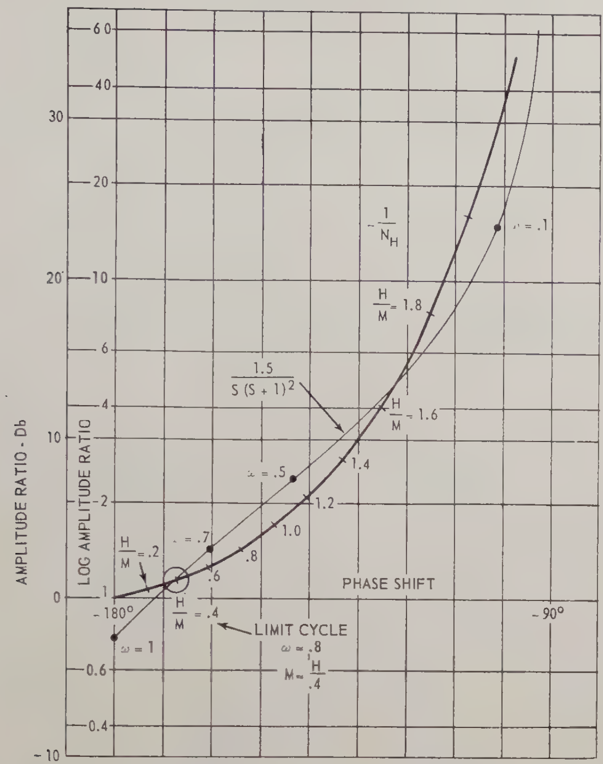

Fig. 7—Amplitude ratio vs phase plot for Appendix I.

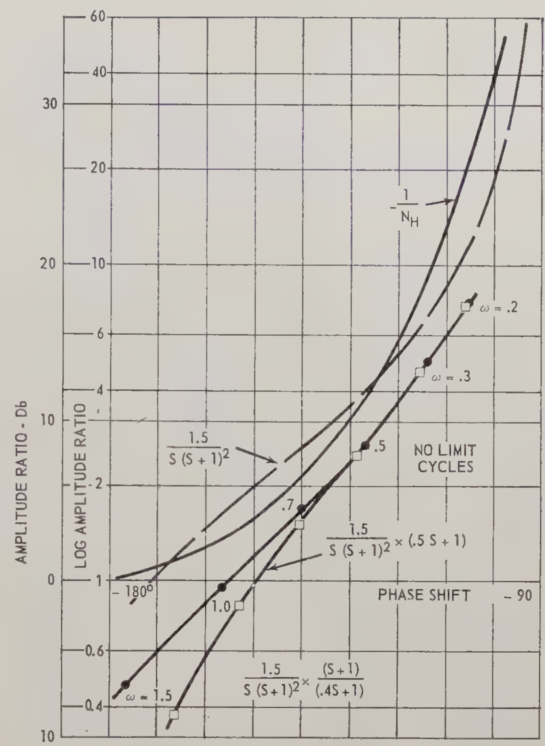

Fig. 8—Lead and lead-lag compensation.

4) Calculate the limit cycle amplitude at the gain element.
5) Plot gain to cause a limit cycle against gain element amplitude and determine the effect of any gain curve by superimposing its characteristic on the same plot.

It will be noted that if the nonlinear elements are separated by linear elements that filter out the higher harmonics, the major problem of the preceding section does not exist. Appendix IV deals with several specific cases of nonlinearities that cannot be combined directly.

Appendix I

Example for Hysteresis Only

Given a system with a block diagram as shown in Fig. 3 and

$$KG(s) = \frac{1.5}{s(s+1)^2},$$

determine the low-amplitude stability of the system.

The equation for the denominator of the system is

$$\left(\frac{1.5}{s(s+1)^2}\right) N_H + 1 = 0$$

from which the function of frequency can be separated from the function of amplitude:

$$\frac{1.5}{j\omega(1+j\omega)^2} = -\frac{1}{N_H}.$$

The two sides of the equation are plotted in Fig. 7 and a limit cycle is found to be present at 0.8 radian per second. Since the linear response to the third harmonic ($\omega = 2.4$) is appreciably lower than for the first harmonic, the analysis is valid.

Appendix II

Examples of Linear Compensation

Given the system of Appendix I which has a limit cycle, determine the effects of various compensating networks on low-amplitude stability.

A rule of thumb for using a lead network in this situation is to position the corner frequency somewhat higher than the limit cycle frequency so the phase shift will be effective without appreciably raising the amplitude. Multiplying $KG(s)$ by $(0.5s+1)$ is found to eliminate the limit cycle as seen in Fig. 8.

Since a lead-lag network has both less phase lead and less amplitude build-up than a simple lead network, the lead term usually should be positioned closer to the limit cycle frequency. The factor $(s+1)/(0.4s+1)$ is shown in Fig. 8 to eliminate the limit cycle.

The system in question contains an integration so lag-lead compensation will probably not eliminate the limit cycle, but should reduce the frequency and the amplitude. Since the additional phase lag from this network must be kept out of the system at the limit cycle frequency, the lag term must have its corner appreciably below that frequency. A network with the characteristic

$$\left(\frac{10s+1}{20s+1}\right)$$

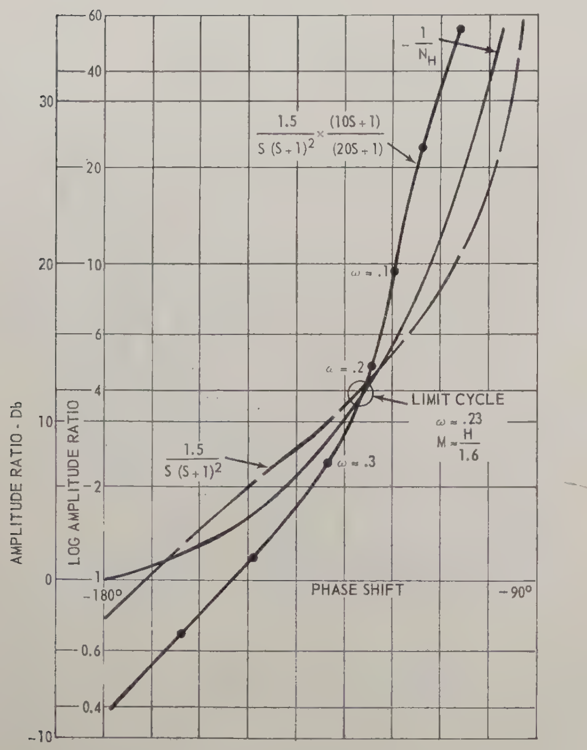

Fig. 9—Lag-lead compensation.

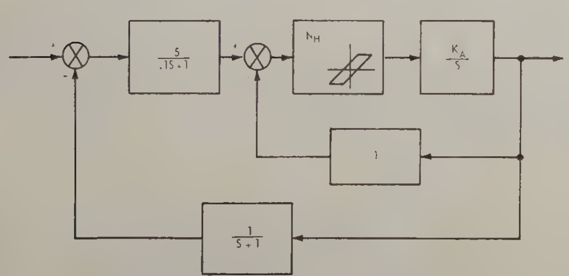

Fig. 10—Block diagram for Appendix III.

reduces the frequency and amplitude of oscillation by a factor of about 4, as seen in Fig. 9. Attention is directed again to the fact that the amplitude reduced is the hysteresis input and the lower frequency may actually cause a higher amplitude in other parts of the system.

Appendix III

Example of Hysteresis and Nonlinear Gain in Direct Series

Given the system of Fig. 10 where K_A/s is an hydraulic valve and actuator combination, determine the effects of a nonlinear K_A on low-amplitude stability. Using the algebra of block diagrams, the critical equation is

$$\left(\frac{5}{(0.1s+1)(s+1)}+1\right)\frac{N_H K_A}{s}+1=0.$$

Rearranging and setting $s=j\omega$ results in

$$\left(\frac{5}{(1+j0.1\omega)(1+j\omega)}+1\right)\frac{K_A}{j\omega}=-\frac{1}{N_H}$$

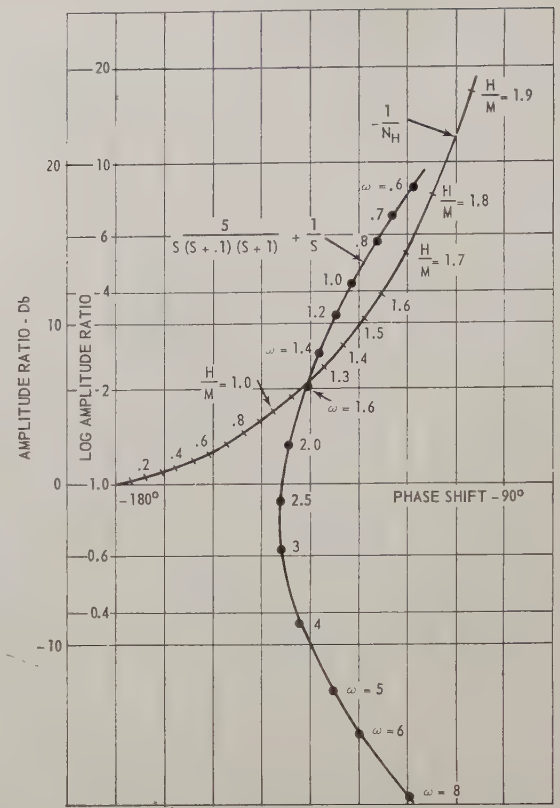

Fig. 11—Amplitude ratio vs Phase plot for Appendix III.

or

$$\left(\frac{5}{j\omega(1+j0.1\omega)(1+j\omega)}+\frac{1}{j\omega}\right)K_A=-\frac{1}{N_H}.$$

The expression in the parentheses can be evaluated at various frequencies by direct combination. With more complex terms, however, it should prove more expedient to evaluate each term with Bode plot templates and make the vectorial addition graphically on a polar plane. Fig. 11 is the combined linear and describing function plot with K_A set equal to unity.

A sample calculation is made now for an ω of 1 radian per second. From Fig. 11, a K_A of 0.73 (−2.7 db) is required to cause a limit cycle which will have an H/M equal to 1.45. Interpolating between 1.4 and 1.5 in Table I gives a hysteresis gain of 0.34 at $H/M=1.45$ so the input amplitude to K_A is

$$A=0.34\times\frac{H}{1.45}=0.23H.$$

When enough points have been determined, several options are available. One method would be to use the approximation for a nonlinear gain describing function and solve for the output of the gain element by multiplying A by the gain. For the $\omega=1$ case,

$$\text{gain output}=0.73\times 0.23H=0.17H$$

Or, if a −6-db limit cycle gain margin is desired, the gain element output as just determined can be divided by 2. When these gain outputs are plotted against the

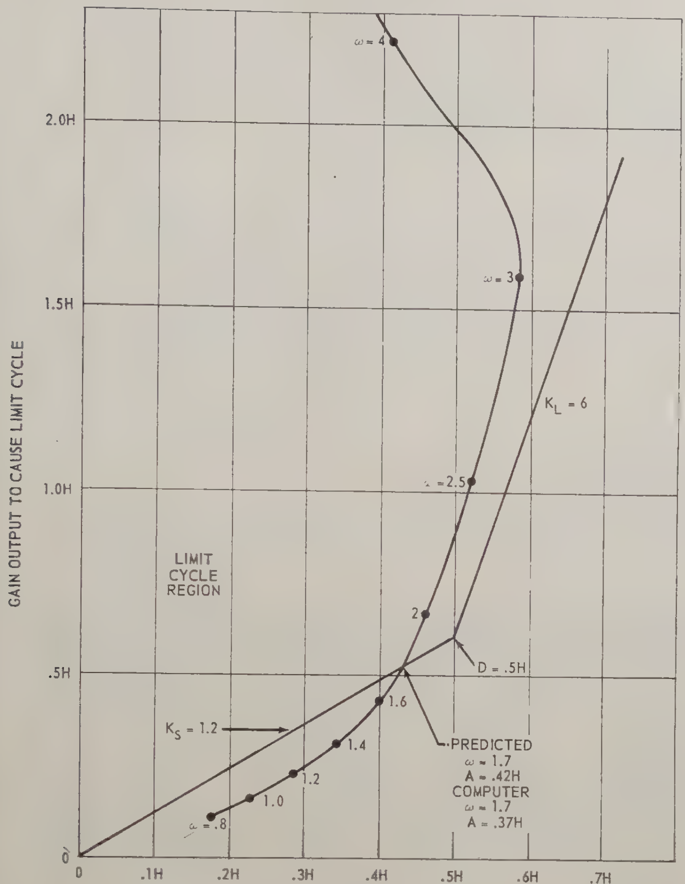

Fig. 12—Gain input for Appendix III.

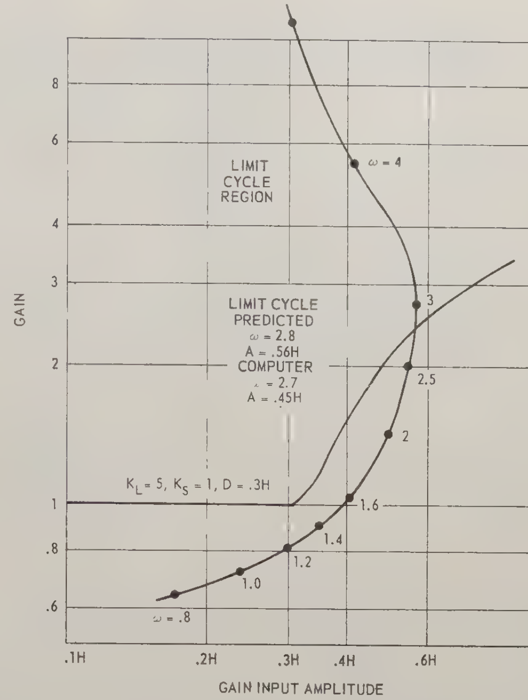

Fig. 13—Gain vs gain input plot for Appendix III.

gain inputs, as in Fig. 12, the effects of a specific gain characteristic (valve flow curve) can be determined or a gain curve can be logically selected, provided that the accuracy of the curve is not relied on too heavily and plenty of margin is allowed at sudden gain changes.

If a gain characteristic as shown in Fig. 5 is dictated by hardware considerations, the log of gain to just cause a limit cycle can be plotted vs the log of gain input amplitude and a set of per unit gain curves prepared to the same scales. The per unit gain curves then are moved about on the limit cycle curve until a satisfactory D, K_L, and K_S/K_L is selected. This method was used for the computer study of the system, where K_S/K_L was set equal to 0.2 and D and K_L varied. Fig. 13 pictures both curves as used for one computer check point. Table II compares predicted and obtained values for a number of cases.

TABLE II

Summary of Computer Results for Appendix III

K_L	K_S	D	Predicted			Computer		
			ω rad/sec	H/M	A	ω rad/sec	H/M	A*
6	1.2	$0.2H$	3.4	1.1	$0.52H$	3.2	1.1	$0.41H$
4	0.8	$0.2H$	2.9	1.0	$0.56H$	2.8	1.1	$0.45H$
3	0.6	$0.2H$	2.4	1.07	$0.54H$	2.2	1.1	$0.40H$
5	1.0	$0.3H$	2.8	1.02	$0.56H$	2.7	1.06	$0.45H$
4	0.8	$0.3H$	2.3	1.07	$0.53H$	2.1	1.06	$0.45H$
5	1.0	$0.4H$	1.6	1.23	$0.40H$	1.5	1.30	$0.26H$
4	0.8	$0.4H$	1.2	1.34	$0.30H$	1.2	1.4	$0.21H$
6	1.2	$0.5H$	1.8	1.16	$0.42H$	1.7	1.26	$0.37H$

* Since A predicted is the first harmonic and A from the computer is the peak value, A predicted should be somewhat larger in every case.

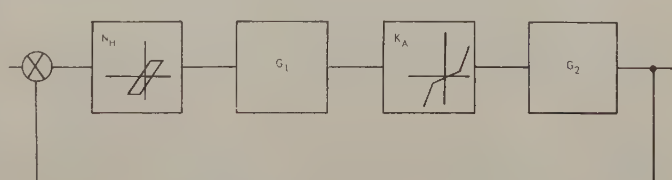

Fig. 14—Block diagram for Appendix IV-A.

Appendix IV

Hysteresis and Nonlinear Gain Not in Direct Series

Part A

To study the small amplitude behavior of a system such as diagrammed in Fig. 14, the critical equation is written

$$N_H \times G_1 \times K_A \times G_2 + 1 = 0$$

from which

$$G_1 G_2 K_A = -\frac{1}{N_H}.$$

The procedure is the same as for the example of Appendix III, except that the input amplitude to the gain is now found by

$$A = M \times \text{hysteresis gain} \times \left| G_1(j\omega) \right|.$$

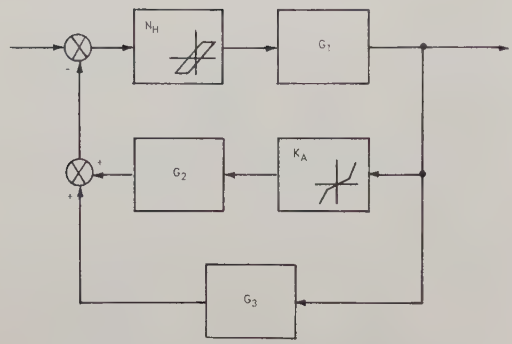

Fig. 15—Block diagram for Appendix IV-B.

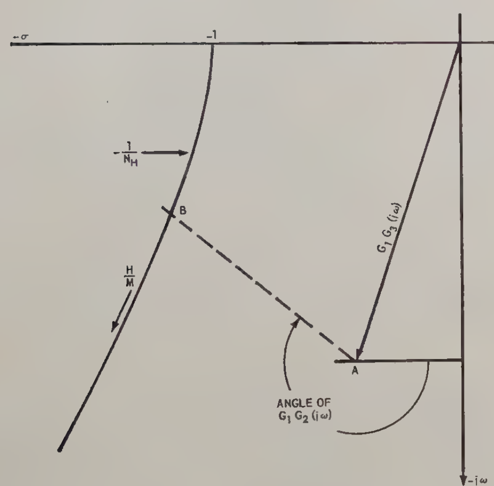

Fig. 16—Graphical construction for Appendix IV-B.

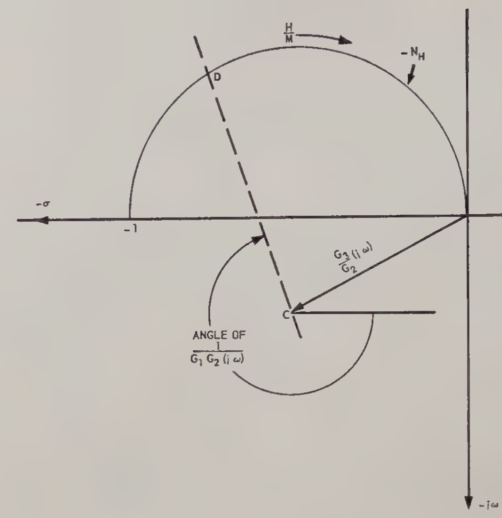

Fig. 17—Graphical construction for Appendix IV-C.

Part B

For a system as depicted in Fig. 15, the equation is written as

$$N_HG_1(K_AG_2 + G_3) = -1$$

and, when N_H is isolated, as

$$K_AG_1G_2 + G_1G_3 = -\frac{1}{N_H}.$$

The solutions to the above equation are found by graphical constructions on a polar plot as shown in Fig. 16. From this the gain to just cause a limit cycle is

$$K_A = \frac{|\overline{AB}|}{|G_1G_2(j\omega)|}$$

and the gain input amplitude is

$$A = M \times \text{hysteresis gain} \times |G_1(j\omega)|.$$

Part C

If the nonlinear gain and the hysteresis are interchanged, the equations are written as

$$K_AG_1(G_2N_H + G_3) = -1;$$

$$G_3 + \frac{1}{K_AG_1} = -G_2N_H;$$

$$\frac{G_3}{G_2} + \frac{1}{K_AG_1G_2} = -N_H.$$

The graphical solution is indicated in Fig. 17, from which the gain input amplitude is found by

$$A = M\,(\text{hysteresis gain} \times G_2 + G_3).$$

Symposium Luncheon

A record of the Symposium would not be complete without mention of the enjoyable luncheon which featured a talk by W. R. Russell, Ramo-Wooldrige Corporation, Los Angeles, Calif.

Although given in a lighter vein than most technical discussions, Mr. Russell's presentation was an excellent introduction to inertial navigation and the problems that attend it. With the aid of colored slides, he described the basic measuring equipment needed, the principles of navigation, the accuracies required in the control equipment, progress that has been made, and the obstacles that are being overcome.

At the end of the luncheon, the second part of the Symposium began.

PART II: OBSTACLES TO PROGRESS IN NONLINEAR CONTROL

PANEL DISCUSSION

A tape recording was made of the entire panel proceedings. Typewritten copies were made from the tapes and sent to each panel member for revision and comment. For the most part, the panel members made few changes, and the personality of each speaker, his subject matter, and his manner of presentation are preserved rather faithfully.—*The Editor.*

Introduction

HAROLD CHESTNUT†, *Moderator*

THE subject of nonlinearity is a very important one. First of all, some nonlinearities are present whether one wants them or not. Second, in many cases by using nonlinearities advantageously, one is able to improve the performance, which might otherwise not be all that could be obtained from the power equipment available. As a matter of fact, in considering the subject of controls, a good deal of the initial work was a consideration of controls from the nonlinear point of view. The paper of Hazen[1] is a classic in this regard. We have been using nonlinear controls for a number of years, using dual-mode type controls, switching from a low-speed system to a high-speed system, and in general, taking advantage of the intuition that we have.

We have been linearizing our systems and trying to solve these nonlinear problems in the linear fashion. It is certainly appropriate that we look to see whether this is really good enough. Now, people who have worked with nonlinear mechanics have brought into the control systems ideas the phase-plane concept. Are they the means whereby we should solve our nonlinear controls? Certainly the describing function is another very useful tool. Are we able to linearize their systems and get out of our problems that way?

Another question which comes up in considering nonlinear control is the fact that we are going to computer control systems more and more, and we're trying to compute an optimum and trying to control to it. Are we going to get ourselves into new problems of stability and accuracy which we did not have when we were merely trying to control to some particular set value? This is an area we'll want to explore and question.

Another problem that comes to mind in connection with nonlinear controls is the fact that occasionally it's desirable to introduce another nonlinearity to compensate for, or to work around, a nonlinear characteristic which is inherent in an element. Can we specify the nonlinearity of a process well enough so that we will be able to compensate for it? Another thought is, will these adaptive servos that we have been hearing quite a bit about, will they be able to perform this optimization for us?

Awhile ago, when the subject of linear controls was getting to be reasonably well understood—I would say about 10 years or so ago—the matter was discussed in a number of circles and John Moore, who is general manager of the Autonetics Division of North American Aviation, came out with a thought that for any linear control, he could always build a nonlinear control which would be better. The question then comes into mind, assuming that this statement is correct, is it worth the trouble to improve a control by purposely introducing nonlinearities?

The purpose in raising these questions is to bring them to our attention in the hope that either our panel members will touch upon them or you in the audience will bring them up later.

Well, with this as a prelude, I would like to open up the formal part of our panel discussion by hearing from Prof. Smith.

† General Electric Co., Schenectady, N. Y.

[1] H. L. Hazen, "Theory of servo mechanisms," *J. Franklin Inst.*, vol. 218, pp. 279–331; September, 1934.

Availability of Necessary Theory for the Analysis and Design of Nonlinear Systems

O. J. M. SMITH†

Introduction by Mr. Chestnut—Prof. Smith is a graduate of Oklahoma A and M, with the B.S. degree in chemistry He later received the B.S. degree in electrical engineering from the University of Oklahoma and the Ph.D. degree, also in electrical engineering, from Stanford University. He is a member of Eta Kappa Nu, Tau Beta Phi, Sigma Xi, and a number of other honorary societies. He's had experience in high voltage, communications, electronic test equipment, and X rays. He's a prolific author and has many technical articles to his credit. He has written a book, "Feedback Control Systems," which is a graduate text on control systems engineering published by McGraw-Hill. He is also a member of AIEE, IRE, American Physical Society, ASEE, and a Fellow of the American Association for the Advancement of Science. At present, he is an associate professor at the University of California on the subject of automatic control and industrial electronics.

In linear analysis, there is a very close relationship between the representation of the system and the representation of the signal. There is also a very close relationship between the synthesis tools and the analysis tools which are used. In general, in linear systems, the transform of a signal which may be some voltage or amplitude, as a function of frequency, and the transform of the impulse response of a system cannot be distinguished one from the other, so that one can interchange blocks at will. You can replace signals by impulse responses of blocks, and this gives you a great deal of flexibility. Also, when you design a linear system and are through with your design, you can analyze it for any other type of signal other than those for which it was designed. For example, a statistically optimum system can also be analyzed. You can say what its response would be to sine waves and impulses. This is not true in nonlinear systems and I think here is where our first difficulty arises when we begin to consider nonlinear systems. I'd like to review a few of the tools that are available and simply state what they're good for and what they can't do.

The first is the describing function method. The describing function can be used to obtain a periodic response of any nonlinear differential equation regardless of the degree of complexity of the linear part of the system as long as the nonlinearity is fairly simple. For example, a two-port nonlinearity can always be handled by the describing function method. One can obtain the frequency of the steady-state oscillations, the amplitude and the wave shape, for a two-port nonlinearity hooked onto any degree of complexity of a linear system. And one can go a little further; one can, for example with nonlinear process control valves define two auxiliary variables, A and B, where A is a function both of frequency and amplitude, and B is a function both of frequency and amplitude. It is possible to obtain a partitioning of the solution of the differential equations into two functions, the first having a gain as a function of A, and the second having a gain as a function of B. Plot these two curves as A and $-1/B$ on log gain coordinates, and from their intersection, find the values of A and B which satisfy the periodic solution of the differential equation of the valve. Then, solving backwards, one may get the amplitude and frequency. In other words, mixed systems with two or three nonlinearities which are dynamic can be normalized with variables which are no longer called just frequencies and just amplitudes, and the values of these variables found quite easily by the describing function method. This gives one periodic responses for nonlinear differential equations. Furthermore, you can design systems like the one shown here today where you use gain vs. phase coordinates. You can design systems that will have no sustained periodic responses. You can design a system that will not have limit cycles. This does not necessarily mean that you have the best system. You've only designed one. It won't oscillate, but how good is it? And this question cannot be answered. You cannot use the describing function method after you've finished the design to say what it would do if you put in step functions, or impulses, or random functions.

So here is a field for further work, which is—to be able to analyze a system designed by the describing function method.

Another type of system is the race-brake system. We had a paper on that topic this morning: a system which uses maximum effort at all possible times until it reaches the final value that it is attempting to achieve. Now, race-brake systems can be designed to follow any sort of an input, like a random input, or an input that is a summation of power curves. These systems generally turn out to have a sequence of transformations of coordinates, a sequence of changes of variables that eventually end up with a decision function entering a saturating amplifier. I'd like to draw a picture on the board of what one looks like. The command is broken into a number of linear computers. We had a special example of this in the final-value problem this morning where this was called an anticipator or predictor and these linear components are followed by nonlinear operations. The feedback is also broken into a series of linear computers and these are followed by nonlinear blocks. The outputs of the linear computers are called the principal axes of the differential equation and Bog-

† Dept. of Elec. Eng., University of California, Berkeley, Calif.

ner is the first one to introduce this concept. Notice that you don't take the difference between the input and the output and operate on the error. You operate on the input independently of the output and then, after the nonlinearities, these are combined into adders and subtractors that weight them appropriately and the output of the whole goes to a high-gain amplifier, followed by the transducer and load. This is a series of transformations of coordinates or change of variables. It performs the same thing as taking a phase space with a switching surface in it and stretching and squeezing up the various coordinates and rotating them until the switching surface is straightened out into cylindrical shape, and then projected on to a plane perpendicular to it, in which case it becomes a single line. Now, the end result of this type of design is that the output will follow an input command closer than any other system with the same restrictions on maximum force. But one cannot, once one has the system designed, turn around backwards and say how this system would operate for arbitrary input signals. It is easy to design the optimum system but it is extremely difficult to analyze it. Here is another place where the field of mathematical techniques for design needs to be extended. Given any sort of an input command, for example, statistical, one can get the best statistical design to accompany the maximum effort transducer, but then one can't say what the system would do for step functions.

A third field is that of random variables passing through two-port nonlinearities. Booton has done some work in this respect. We can say what the power spectrum, the autocorrelation function, and the probability distribution of a signal will be after it goes through a nonlinearity, such as a clipper or a dead zone or a power-law nonlinearity. This is easy to do if the input is Gaussian, but is extremely difficult to close the loop around this nonlinearity and then say what the spectrum is in the input and output. Here is another field that requires considerable work. Now there is a nice approximation. If one filters the output of the nonlinearity sufficiently and takes out all the harmonics that are generated, then, like the describing function method, one can have a quick approximation by finding an equivalent linear gain which is a function of the fundamental component. (Gaussian distribution.) That is, that component of the output which correlates with the input. This equivalent gain of the nonlinearity to a random Gaussian signal can be used in complex systems. But in simple systems it cannot be used unless there is a considerable amount of harmonic filtering.

A fourth type of nonlinear system is a modulation system. Given that the signal spectra are very much lower in frequency than the carrier frequency, it is fairly easy to assume that modulating this spectra on top of a carrier, say 400 cycles, operating on it with various networks and amplifiers, and demodulating with a two-phase motor, the signal which comes out is essentially equal to the one put in except for those deliberately introduced operations, such as phase lead that might have been introduced. This is true until the signal frequency begins to approach that of the carrier frequency. When the two are the same order of magnitude, the computation of what comes through is not too difficult as long as it is an open-loop system. But, as soon as the loop is closed, all the second, third and fourth harmonics that are generated by the demodulator come back to the input and are remodulated again. Some of these higher harmonics beat back down into the signal range, and it becomes quite a mess. Here is a field for a mathematician to work out, and I think this is a fairly straight forward problem—to work out the exact analytical solution for a modulator, filter, demodulator, and filter in a closed-loop system.

Last, I'd like to mention the phase-space or phase-plane technique. The phase plane is very excellently suited to analyzing second-order systems. If one only has a second-order system, this is fine. It has the advantage that one can consider a very complex nonlinearity with a very simple linear or dynamic part. The describing function method, on the other hand, has the advantage that with it you can consider a complex linear system of very high order and a relatively simple nonlinear part. The phase space, it seems to me, is useful to a man whose mind works in multidimensional space, namely, a mathematician; and this can be made useful to engineers only if these multidimensions are brought down into multiple voltages on wires where one can put his fingers on them. For this, then, I would recommend that phase-space techniques will be of value if they are used as a tool for developing changes of variables that can be understood by the practical engineer concerned with building a network for controlling nonlinear systems, particularly those of the race-brake or maximum effort type. I hope I've given you just a very brief cut across the tools that are currently being used and my opinions concerning their usefulness.

Nonlinearities in Machine Tools and Missiles

JOHN L. BOWER†

Introduction by Mr. Chestnut—Our next speaker will be John Bower. He received the B.S. degree from the U. S. Military Academy and the Ph.D. degree from Yale University. He is a member of the IRE Papers Review Committee; a member of AIEE, and a member of Sigma Xi. His experience includes that of being an associate professor at Yale; he worked at General Electric in control instruments. He has done fire control and flight control work at North American. He is the author of a number of articles, including those on amplidynes, resistance capacitance network synthesis, machine tool control, and is a contributor to the book on "Automation in Business and Industry," which Gene Grabbe edited and which was published by John Wiley. Currently, he is a section chief of the Automatic Industrial Controls Section in the Autonetics Division of North American Aviation.

I MIGHT SAY in the beginning that there is, after all, some connection between inertial guidance and the control of machine tools, when you consider that the problem of machine tool control is to guide a metal cutting instrument along a path on the surface of a part. The difference appears in two striking ways. First of all, in machine control you are dealing in terms of thousandths of an inch instead of miles, which is a difference of about eight orders of magnitude, and in the second place, you are dealing with kilobucks instead of megabucks. I can perhaps point out a few similarities and show the carryover of some of the theory. In a typical schematic representation of a machine tool, five axes of motion are needed: three cartesian axes and two angular degrees of freedom. I point out here that the use of a gimbal system, somewhat similar to what you have in guidance systems, is convenient in machine control as well. But more than that, if you study problems related to the motion of a machine of this kind on the surface of a part being cut, you find that the problems of the representation of the motion and the effective error in the system is most conveniently handled in terms of vector language. The problem that arises in moving the machine, likewise, is in many cases a vector problem, and I would like to suggest the possibility of applying that concept in some of the other areas that we may discuss this afternoon. The vector language is useful where there are many simultaneous errors that need be considered. An error, for instance, in the X axis, is not equivalent and cannot be summed with an error in the X axis, the Z axis or either of the two angles to determine the quality of the final result. Instead, it is necessary to consider the vector error.

Now to touch upon another point, I'll simply mention that the solution of problems in the numerical control of machine tools does not appear to demand very much use of sampling theory. Sampling theory can be ignored simply because the time scale of the pulsing system is much shorter than the time scale involved in the mechanical devices that are used. By the way of explanation, I might say that the loop used in a great many systems consists, first of all, of a measuring device that delivers a pulse representation of the motion of the machine into an error accumulator or register. Simultaneously, a pulse train defining the reference motion is fed into the same accumulator, but in the opposite sense. The accumulator consequently delivers out the instantaneous error which is taken through a digital-to-analog converter to provide an analog error signal to an amplifier controlling the motion of the machine.

As I said, in the generation of the pulses, the time scale is sufficiently short that we do not need to give much consideration to the sampling problems. However, there is one digital aspect that does have to be taken into account; that results from the fact that the operation of the error accumulator is digital and delivers an error signal which is, you might say, a staircase function of the true error. The output rises only in discreet steps, as the error is gradually increased, up to the point where the accumulator spills over. This gives rise to a problem that, on a small scale, is like the "bang-bang" or relay servo, and it also could be a problem depending upon just how the system is designed. Now, it is possible to design systems that avoid this situation, or alternately to reduce the least count of the gauge to the point where the staircase nonlinearities are of little importance.

Unfortunately, not all of the problems that arise in this field can be handled by the kind of compromise I have just been talking about. There are some real ones that do face us and are not completely solved as yet. As an example, at several points in the system there is a significant amount of stiction in combination with elastance. If, we drive the machine table by means of a lead screw or with any type of actuator that is less than mathematically perfect, we will have a good deal of stiction present. This is particularly true if we substitute a piston for the lead screw and have a much more resilient oil column to drive the machine. However, even in the cases where a large steel screw is used in driving the machine, there is no way at present to completely get rid of the elastance and the stiction problem. The velocity of a compression wave in steel is such that at 120 cy-

† North American Aviation, Inc., Downey, Calif.

cles per second, for example, the compression wave will go through a phase change of about 60 degrees in 20 feet; that is, in a uniform steel column. Now you can imagine what happens when you load a couple of tons of metal on a screw that weighs one fortieth as much. In this area, we seem to be up against a problem that has no immediate simple solution since we find that the cutting element operating on the work piece can very easily generate forces of several thousand pounds at frequencies in the vicinity of 100 cycles per second and, in order to get good surface finish, we are required to hold the part in position steadily. We seem to have arrived at the type of problem that the electronics industry encountered some twenty years ago, namely, the transport lag and the fundamental problem of spacial distribution in our control elements. This problem is one that is going to require as much attention from the engineers who contribute ingenious design techniques as it is from those who provide the mathematical tools for the solution.

Nonlinearity in Process Systems

ERNEST G. HOLZMANN†

Introduction by Mr. Chestnut—This time we'd like to branch off to a little different field of activity, namely that of process control. The next speaker will be Ernest Holzmann, who has received the M.S. degree in electrical engineering from M.I.T. He worked as a process control engineer with the Shell Development Company, Emeryville, Calif., before he joined General Electric's Atomic Power Equipment Department, where he is currently employed as a process dynamics engineer. He has written a number of papers on the subject of process control.

Introduction

THE needs of technology frequently force engineers to find practical answers to problems which, at the time, seem to defy any known theoretical approach. The control of nonlinear processes is no exception.

Control engineers in the chemical process industries have long recognized the possibility of achieving superior responses by means of nonlinear control. To cite just one example: the characterized valve introduced by Dr. C. E. Mason about 25 years ago was a deliberately nonlinear component. It has since come into almost universal use. Nonlinear controllers, too, have been applied to the process industries for years.

Despite these very real accomplishments, process control engineers would be the first to admit that neither the theory nor the hardware of nonlinear control have progressed very far in the time since Dr. Mason's invention.

Three obstacles stand in the way of progress: 1) inadequacy of mathematical training; 2) lack of precise knowledge concerning the dynamic and other characteristics of the processes to be controlled, and 3) economic factors.

† General Electric Co., San Jose, Calif.

Mathematical Training

Nonlinear techniques require specialized mathematical training. Recognizing the need, more and more colleges and universities are now teaching courses in feedback control theory. They usually include phase-plane analysis and quasilinearization techniques.

The lack of adequate training should not, however, be a permanent handicap to any engineer. There is no shortage of books and technical papers dealing with nonlinear differential equations. Probably everyone of us here feels that he could benefit from further study of the available material.

If all else fails, the engineer can resort to the analog or digital computer to solve his problem. You will have noticed that nearly all technical papers concerning practical aspects of nonlinear control base their conclusions on results obtained with an analog computer.

System Characteristics

There is no way of bypassing the second obstacle. You just cannot hope to improve conventional control schemes unless you learn more about the processes you are trying to control. The required information may be developed by tests or by analysis.

Testing

Several of the larger companies have initiated programs of dynamic testing of process equipment both in the laboratory and in the field. Although very costly, the price paid for test equipment and manpower is small compared to the savings that can be made by improved control methods suggested in the course of the tests. Sometimes these savings are very impressive and pay handsomely for the time and effort spent in testing. Be-

sides solving the immediate problem, such tests sharpen our understanding of process dynamics. Thus, they acquire added significance and become potentially valuable not just to one particular company, but to the entire industry.

Analysis

In parallel with the testing of process plants, increasing effort is being expended on the analysis of the basic transient phenomena of mass and energy transfer. Here, it seems that the more we learn, the less we know. The goal is to develop mathematical models that behave like the physical process. What makes progress so particularly difficult is that we are usually dealing with nonlinear, distributed systems, some of whose parameters vary with time.

Self-Optimalizing Controller

The only type of control which does not require previous exact knowledge of the process is the technique of self-optimalization. This technique may lead to a controller which is simpler to adjust than conventional proportional, reset, and rate action controllers. While the self-optimalizing controller holds great promise, its engineering development has barely begun.

Dual-Mode Controller

The dual-mode controller is more likely to prove practical in the near future. The term "dual-mode" refers to the proportional gain which increases sharply with error. The advantages of the dual-mode controller in servosystems are well known. What is not so well known is that the same advantages may not be realizable in the control of industrial process plants. To be effective, a dual-mode controller must operate a final control element (normally a valve) that is free to move rapidly through its stroke length. Such operation can, however, induce transients elsewhere in the plant. The result may be worse than the original disturbance responsible for the control action. In this event, nothing is gained by departing from conventional control.

Very often, poor control is due to transport lags in the process or in the sampling line between the process and the point of measurement. If the lags are irreducible, a cascade arrangement of conventional controllers may relieve the problem far more effectively than dual-mode control.

Undoubtedly, the dual-mode controller has a place in process control. Much research, however, needs to be done to determine what kinds of process would benefit from it.

Economic Factors

Any technical advantage of nonlinear process control must be weighed against the additional cost. Factors to be considered include:

1) The man-hours required to develop specifications for a nonlinear controller in any given application, as against specifying a linear controller.
2) The design and manufacture of a nonlinear controller having the required degree of flexibility and reliability, as compared with standard linear controllers.
3) The training of operators who can adjust a nonlinear controller so that its potential for better control is actually realized.

So far, there is no evidence of definite economic advantages to be derived from nonlinear control. The required evidence may be brought to light through applying operations research methods to the whole problem of the cost of process control.

In the meantime, however, neither users nor producers of control equipment are inclined to channel large sums of money into the development of nonlinear techniques or hardware. Until they do, progress in nonlinear process control will continue to be slow. It will depend on companies who take an attitude in advance of the most economical. They are willing to sacrifice a margin of profit for the sake of promoting the art and the science of control. By so doing, they make a contribution no less important than the technical accomplishments of individual workers.

Conclusions

Progress in nonlinear process control has been relatively slow, despite the significant work of some individuals and companies. Better training and more extensive testing and analytical work will result in new applications of nonlinear control. The present rate of progress could be accelerated if the economic advantages of nonlinear control were clearly visible.

The Role of Computers in Analysis and Design of Control Systems

GEORGE P. WEST†

Introduction by Mr. Chestnut—Our next speaker will be George West. George received the B.S. degree in electrical engineering from the University of California, and then put a stint in with the U. S. Army Signal Corps as a radar officer. Following that, he received the M.S. degree in electrical engineering from Stanford University. After the war, he worked for NACA at Moffett Field, Calif., in the area of research and development. Later, he went to General Electric and served as a supervisor of computing groups operating both analog and digital computing equipment. Presently, he is working at Ramo-Wooldridge in the Digital Computing Center.

THE role of computers in analysis of control systems seems to be rather straightforward and perhaps routine. This impression is due to the publicity that has accompanied each new computational accomplishment. As a result, the feeling is prevalent that any problem can be easily solved with the aid of modern computing equipment. While many problems of control system analysis are profitably solved by modern computation equipment, there are control systems which are very difficult to analyze with the most powerful of the modern computing equipment. These difficult problems of analysis could, of course, be solved with modern equipment, if cost, or elapsed time, were not serious considerations, by simply letting the computer grind long enough. If the control system analysis produces a system of nonlinear differential equations of moderate order with a response which contains both high and low-frequency effects, the analysis may be difficult particularly if the designer is interested in both of these effects. If an analog computer is employed for the analysis, the system of equations will often require more nonlinear elements for its solution than are available. A system of moderate order may require more multipliers than are available on many of the large analog computers. If a digital solution is employed, the numerical integrations will require a small time step to faithfully reproduce the high frequency effects. If the low-frequency effects are to be observed, the integrations must be carried forward over a great many of these small steps. The digital computation will require an excessive amount of computation time. Problems of this type are very profitably attacked by a combined analog-digital attack.

The Ramo-Wooldridge Corporation employs a system which consists of an analog computer (approximately 300 functional amplifiers) and a digital computer (ERA 1103A) communicating by means of an electronic analog-digital conversion system (15 A/D channels and 15 D/A channels with channel selection and control). The problem is separated so that low-frequency loops are solved digitally while the high-frequency loops are solved on the analog computer. Usually there are loops which can be shifted to either machine. By shifting some of these tasks from one machine to the other, it is usually possible to balance the load between the two machines. Another interesting application of the analog-digital approach uses both machines to perform a single computation, with one correcting the other. With this technique the analog computation follows the high-frequency variations while the digital computation periodically corrects out the low-frequency drifts.

With the exception of the small class of problems discussed above, which may require a combined analog-digital approach, the role of computers in the analysis of control systems is well established. Computers provide a ready means of analysis in control system studies. The role of computers in the design of control systems has not been exploited. In general, any design or synthesis problem is more difficult than the corresponding analysis problem. Of course, analysis can be used as an aid to design. In this approach the systems engineer is designing by a cut and try process. Instead of building system hardware and modifying it, he conducts his experiment on a computational model of the system. The computational model is usually much cheaper to modify than actual hardware.

Just as one would hope to find a better design procedure than cut and try, it is hoped that a better use of the computer can be found for control system design. One approach is to let the computer conduct a series of experiments and modify its own response so as to find some sort of an optimum. Initially these approaches will be little better than the cut and try approach outlined above. However, as improved techniques of specifying desired response and of measuring deviations from the desired response are evolved, this method will become more useful. Dr. Bellman's dynamic programming[1] seems to offer a new computer approach to control system synthesis.

In summary, computers are widely used in the analysis of control systems. Computer applications to control system design are presently implementations of the cut and try approach. The future will probably see more sophisticated applications, however.

† The Ramo-Wooldridge Corp., Los Angeles, Calif.

[1] R. Bellman, "Dynamic Programming," Princeton University Press, Princeton, N. J.; 1957.

Problems of Nonlinearity in Adaptive or Self-Optimizing Systems

CHARLES F. TAYLOR†

Introduction by Mr. Chestnut—Our next speaker is Charles Taylor. Dr. Taylor went to the University of Massachusetts and received the B.S. degree in electrical engineering. He received the M.S. and Ph.D. degrees in electrical engineering at Stanford University. He is a member of Sigma Xi and the ISA. Dr. Taylor has worked extensively on nonlinear mechanics and systems control theory. At the Air Force Cambridge Research Center, he investigated nonlinear techniques for sampled-data systems as applied to track while scan radar and weapon control computers. At Varian Associates, he worked on pulse-circuit design and theoretical electromagnetics. At Daystrom, where he is currently located, Dr. Taylor is Head of the Systems Analysis Group.

WELL, this is a bit simple-minded, but to set the stage, let's describe the general control problem. We have a process to be controlled. We have sensors or transducers with which we gather information about the process and we have devices that digest or operate upon the sensed process information to produce control signals which in turn become muscles or actuators that adjust process variables. With this basic viewpoint of control, one of the first hard facts of life that we run into is that the processes we wish to control are often nonlinear and/or time-varying and thereby defy convenient analytical description. Another important hard fact of life is that the sensors with which we view the process are not error free. This means that there is often a discrepancy between the actual state of a process and the observed state of a process.

I believe that these two (as I have called them, hard facts of life) coupled with the continued desire to improve process performance are the real basis for nonlinear adaptive control systems. In fact, these considerations may be used to define adaptive control as follows:

> *Adaptive control is a method of control aimed at obtaining optimum system performance even when there exists incomplete or inexact analytical or analog model of the process that is being controlled.*

At this point, with a working definition of adaptive control, we come to the main purpose of this talk, namely, to discuss some of the approaches to adaptive control and the problem or lack of problems associated with these approaches.

The now classic approach is the Draper-Li method of optimalizing control where a continuous "hunt and seek the optimum" game is played on the process, *i.e.*, certain process variables are perturbed or varied in controlled fashion in order to extremize a specified output variable. The controller seeks the peak of the performance curve. The major problems or limitations with this approach are:

1) Time required to locate the optimum becomes exceedingly long if the process is complex from the viewpoint of multivariable and/or multiple time lag.
2) Degree hunting of the input process variables necessary in this approach is not acceptable in a large number of industrial processes since it can lead to large scale process upsets (instability).

Zieboltz and Paynter have suggested the use of a fast-time analog of the process to overcome the main limitations of the Draper-Li approach. Here the "hunt and seek the optimum" game is played on the fast-time process analog. Only the optimum control settings are applied to the actual process. The main problem with this method is to obtain a good fast-time process analog under all operating conditions. If this were possible in the first place, it would mean that the process could be described analytically. This is seldom the case.

A more realistic approach to adaptive control has recently been suggested by Eckman and Lefkowitz. The idea is to start with an imperfect analog of the process (either real or fast time) and to continuously or periodically automatically modify or refine the model to better agree with the actual process. These investigators call this approach self-checking. Needless to say, all optimizing is done using the analog of the process rather than the process itself.

I believe that the Eckman-Lefkowitz approach suffers from one disadvantage that is inherent in the other previously mentioned approaches. This is that optimization is conceived achievable in a steady-state sense. By this I mean that the controlled experimentation of cause and effect perturbations (in this case on an analog of the process) in seeking optimum performance are carried out neglecting the process dynamics. No attempt of optimizing control is made during transient or upset conditions.

Finally, let me brush on what I feel are two general problems or obstacles in nonlinear control. One is the lack of useful existing mathematics or mathematical tools. Here I feel that there will be required a bootstrap-type operation between mathematicians and operating control engineers. The common meeting ground may well be computer experimentation. The other obstacle is the tradition-mindedness of all of us in the control field. For example, we like to go on thinking in terms of linear models even when we are past the boundary of linearity. I feel that this attempt to preserve the conveniences of the linear world (the principle of superposition, for example) tends to suppress independent advancements in both mathematical and physical aspects of nonlinear control.

† Daystrom, Inc., La Jolla, Calif.

Open Panel Discussion

ALTHOUGH the discussions between the panel members and the audience were tape recorded, they are not reproduced here in detail because the topics and speakers changed frequently, the identity of the voices was sometimes difficult to ascertain, and the individual statements were not always recorded intelligibly from the floor.

However, the principles of the arguments were obtained and the basic ideas involved are preserved for further meditation in the following summary which is tinctured with exact quotations as the thread of the actual discourse is followed.

The first discussion originated with a question from the floor concerning self-adaptive controls. In particular, the panel members were asked if they could provide more information about a device in which a parameter was not varied to improve control but where a simple model of the desired system, operating in parallel with the actual system, was used to determine the desired output. Then, this output, and possibly its higher derivatives, was compared with the actual output to provide a correction signal to the actual system.

Dr. Taylor was not familiar with this system, but he questioned the benefits of feeding an error signal back through a high-gain amplifier. Others in the audience had divided opinions as to whether the high gain needed to provide the necessary correction would be detrimental or not. However, one participant claiming to have constructed such a system three years ago stated that a large gain was necessary for satisfactory operation and that this caused stability problems.

Another person compared this system with the "conditional feedback" system described by Lang and Ham,[1] known in other similar forms as a feedforward system or as an open cycle-closed cycle system. Mention of conditional feedback induced Prof. Smith to end the discussion with the comment that " . . . This is essentially a feedforward system and if it works at all you are throwing away your feedback system and using just the feedforward block. The one that Lang and Ham had is a balanced bridge and is extremely sensitive to component variations.

The second topic from the floor also concerned adaptive control of the brain and nervous system. A description was given of neurophysiological experiments on cats. Electrodes were inserted into a cat's nervous system along an auditory pathway. Audible clicks were presented to the cat and the electrical activity at the electrodes was observed. After a few treatments, the magnitude of the electrical signals would decrease at each of the electrodes, which were essentially detecting signals in anatomical feedback loop. However, after this habituation or conditioning to sound, the electrical signals could be increased again if the cats were given electrical shocks simultaneously with the audible clicks.

This example was used to illustrate an actual nonlinear control system where unimportant input signals can be reduced to a level of no response, but not below a monitoring level. This is similar to reducing one input to allow the reception of more important signals until that input again becomes important.

There was further speculation that signal adaptive systems might find use in process control or other electromechanical systems, and the subject was extended by the suggestion that signal adaption would be useful to a creative engineer who must learn how to ignore extraneous inputs and determine what factors are important to his problems. From this comment a new topic of discussion emerged concerning the training of engineers and the proportion that should be maintained between pure and applied mathematics in an engineering curriculum.

The first argument was that pure mathematics can actually hinder engineering progress if an excellent solution to a practical engineering problem is overlooked because it is not mathematically rigorous or if an engineer becomes so involved in the mathematics that he never arrives at a practical solution. The describing function was cited as an example of a practical engineering tool that is not accepted by mathematicians as a solution to a nonlinear control loop problem. Basically, the question raised concerned the problem of liaison between mathematicians and engineers, the problem of obtaining interpretations of developments in pure mathematics and transforming them into useful engineering techniques. Should this be the function of a mathematician or an engineer?

Mr. Holzmann made the first reply, "Just one comment on this general problem. It has been considered that what we need is handbook information on the mathematical tools that are available to solve our problems. Now, certainly that would be very useful to have. But it would only be useful to those people who rubbed elbows with these techniques before. Take, for example, the technique of phase-plane analysis. It isn't necessary that you understand all the mathematical foundations of the phase-plane approach to the solution of nonlinear differential equations in order to solve a control problem. But it is necessary that you know what kind of problem you can solve by this technique, and then you can go to a handbook. If a man has never heard of the technique, or having heard of it has never used it, I am afraid the handbook wouldn't be much good to him. He would just pass over those sections of the handbook."

A member of the audience answered that the obstacles to nonlinear control should be thought of in two different ways. One of techniques and the other of theory.

[1] G. Lang and J. M. Ham, "Conditional feedback systems—a new approach to feedback control," *Trans. AIEE* (*Applications and Indus.*), no. 19, pp. 152–161; July, 1955.

He pointed out that techniques might be presented in handbooks, but that handbooks, like machines, cannot be a substitute for the brain, and that perhaps it would be better if engineers became more like mathematicians.

Mr. Chestnut diverted the conversation slightly by asking Gene Grabbe, Chairman of the PGAC, whether the PGAC had any future plans for making tutorial papers available to engineers to keep them abreast of recent mathematical and engineering developments.

Dr. Grabbe replied that the PGAC is a young group, that it has been trying to do this on the local level at chapter meetings, but that the greatest progress will probably come from a higher level of cooperation between the various societies. He pointed out that the PGAC and the AIEE and ASME Control Committees are working toward this goal. He also noted that if tutorial papers at meetings or in the TRANSACTIONS should be increased to meet the demand, any suggestion would be welcome because the real purpose of the Group is to satisfy the needs of its members.

Then Dr. Grabbe changed the subject by noting that business uses two thirds of all digital computers produced for economic reasons, and he asked the panel to comment on optimizing the profit of a system such as those used in the chemical processing field.

Dr. Taylor remarked that at the present time this was done by experience or by averaging economic data obtained in the past. Although this was a necessary step, there is a definite need to operate on present data as well.

Mr. Holzmann supplemented this comment with a broader statement: " . . . Gene has brought up a very interesting topic. In connection with the process industries, nonlinear control has two distinct areas of application. One is nonlinear *control of process variables* by means of analog or digital type controllers. The other one is *profit control*. Profit is a highly nonlinear function of the set point values of the process variables. Someone else has already defined what a set point is in a chemical plant. So it is generally understood what I am talking about. It certainly has very little to do with the dynamic response; it has a little more to do with the way different variables are interrelated in the steady state. We are assuming, in other words, that a steady state has been established one way or another, and we are now interested in the profit function as a function of the different variables. For instance, as a function of purity; as a function of the different raw materials that you could conceivably use as alternatives in feeding your plant, according to the market situation, the distribution of your refineries across the country, and the transportation problem.

There are many problems involved here which rightfully belong, I understand, in the area of operations research. It's a pity that no operations research man is here on the panel to present the views of that wide and increasingly important field."

There were further miscellaneous comments that the economic factors occur on a different time scale from the closed-loop control, and that the mathematics needed in solving profit problems may involve nonlinear programming, a field which is not well developed.

The subject of mathematics was resurrected again by comments to the effect that mathematicians should not become involved in engineering because a mathematical problem is often solved when it is found that a certain solution is possible, with a certain form, satisfying particular requirements, and these are of little significance in engineering.

Another rebuttal defended the use of analog computers, not as a substitute for the brain, but as a tool to determine whether higher order effects derange the desired system performance, and also as a tool to vary system parameters, a feature not susceptible to formal mathematical processes. These comments ended with a request that Prof. Smith illuminate this problem of system optimization by enlarging on some of the techniques that he described briefly in his talk on availability of theory.

Prof. Smith complied with the following discourse: " . . . The method I would like to enlarge upon is the technique of following the lead of mathematics, since what we have just been talking about has been mathematical tools. This occurs, in one case, when one is solving a system using the maximum possible effort on the output. If one sets up differential equations for a system using the force member with maximum force in either of the two possible directions, or zero force in the middle range, and makes the output variable match the input variable exactly after some specified time, one can solve these equations going backwards in time and end up with a set of differential equations that fit together with initial conditions of one being the final conditions of the previous. And when one does this, one has a whole collection of linear functions and a whole collection of nonlinear functions. Using the guide of the mathematics, I would say that every time you get a pocket full of linear functions, call it some new variable. Every time you get a pocket full of nonlinear functions, call it another new variable. In doing this, you can eliminate these variables and you will end up then with a Christmas tree of variables, where you generate one from the previous by means of some stated operation. The operation is that basketful that you found on the right-hand side of the equals sign. Now, I think this is an extremely fruitful technique to follow. Don't try to force the mathematics to take your preconceived form; instead, you follow along like taking a canoe through the rapids, and when you are done, you'll have a great big block diagram and, in general, these block diagrams can be greatly simplified once you get them. I would like to say that I would recommend, in following the mathematics of a new-type problem, that you obtain your intuitive clues from the way which it seems would best simplify your mathematical equation."

Prof. Smith had scarcely finished when Dr. Taylor

said, "I just have one point of disagreement with Prof. Smith. I do believe in your approach. You already have forced the mathematics to follow where you'd like to go because you only consider bang-bang systems and this is not an autonomous case. This can be reduced to the nonautonomous case and fit space-plane techniques. However, I really do disagree that it certainly may be a fruitful approach as a guide, because I don't think that it is the only way, and because I have yet to find any real leads, mathematically speaking, towards a non-autonomous case where we have other inputs that can be reduced to essentially no inputs by transformation of variables where we are back into phase-plane techniques."

There was a compromising statement from the floor indicating that although there was not a great amount of mathematical information available, there has been some work on defining explicit bounds on the behavior of a nonlinear, nonautonomous control system in the presence of an arbitrary input if certain bounds in the derivatives and magnitudes of the inputs were available. It was noted that the principal difficulty was the greater range of problems that were encountered in the nonautonomous system than in the autonomous system, and that although mathematical techniques are available they are untried because of the difficulty in applying them.

The discussion was directed into other channels by Mr. Chestnut, who observed that mathematics is probably not the sole obstacle to progress in nonlinear control, and that other areas should be considered.

The next area to be considered was that of time-sharing signals in a control system. The problem was discussed whether it is necessary to have signals available at all times, and if not, how frequently should they be present. For example, a single computer can be used for different lengths of time to operate on signals in different parts of the system. This would involve a combination of analog and digital mechanization and analysis.

Dr. Taylor was glad to contribute some thoughts to the discussion by noting that basically this problem of sampling in mixed systems was typical of process control where minor loops are analog and the over-all control loop is closed by a digital controller which adjusts set points. He pointed out that this is essentially a sampled-data system especially when the inputs, such as those obtained from flow-meter paddle wheels, are often obtained in discrete form.

Mr. West was asked whether computers would be applied to solve the problem, and he replied that if a particular system were specified this could be done, but as an aid to analysis or synthesis of a mixed control system the computer would have to be used on a trial or error basis.

The discourse on digital systems was turned to Dr. Bower when he was asked to comment on the feasibility of applying the Inchworm actuator to machine tool control.

Dr. Bower gave his version of the Inchworm as a device " . . . which uses a magneto-strictive element that is clamped to an actuating rod on one half-cycle, after which it is changed in length by magneto striction. It is then unclamped and restored to original length in the next half cycle. Naturally, because it is a magneto striction device, its typical length extension is very short, and a reasonable rate of motion is obtained by reason of using very high frequencies." He revealed that he had not had experience with the device, but that it would undoubtedly have important applications in many cases because of the inherent stiffness of the magneto-striction drive. However, he pointed out that hydraulic drives could be substituted for it in many instances with satisfactory results.

The discussion returned to the obstacle of education, or lack of proper engineering education, for engineers and technicians. A question was asked that if it is often difficult to educate technicians and some engineers about the concepts of linear system operation, how will it be possible to give an understanding of nonlinear system characteristics?

Mr. Holzmann was asked to comment on this, and he said, "Well, you have already seen how we, as engineers, try to deceive ourselves into believing that a nonlinear block, in a block diagram, behaves like a linear one. It is my conviction that we have to do the same thing when it comes to installing that nonlinear block in a plant. We have to pretend, as far as the operator is concerned, that he has the same sort of thing there that he has always dealt with when he had a linear controller, for example, the two-mode controller. I see no reason at all why the operator should have greater, or let's say much greater difficulty, with that than he is presently having with his reset, rate action, and proportional-type controller.

"Perhaps we can do away with one of these functions and replace it by the break point, the knee, in the nonlinear characteristic; so there won't be any more knobs, there will just be new rules of thumb."

A reminder was injected by Mr. West that the operators mentioned by Mr. Holzmann are skilled operators who have been running plants for 25 years or more, and that if their experience could be understood and programmed, automatic process control would be nearer reality.

This point was further emphasized by a speaker from the floor, who repeated an anecdote he had heard " . . . that a very difficult missile problem is to take an oblate spheroid, transmitted from a moving target through a moving atmosphere and hit it, or having transmitted it from a moving platform, hurl it through a moving atmosphere to a moving target. . . . The computer that you build to do this would certainly fill this room if you did a good job. (Yet) you can take a half-witted halfback and do it nine times out of ten any Saturday afternoon all fall. (The point is) that last year, some place, this happened to many of these operators of

much better computers than you have in your engineering departments. . . . "

The topic of discussion was diverted again to Dr. Bower, who was asked whether the numerical control problems in machine design were being solved by machine tool companies or by electronics companies. Dr. Bower reported that the machine tool companies have a much better understanding of the structural and mechanical drive problems, and that electronic companies have developed the servocontrols. However, he noted that the machine designers were becoming more involved in the design of the servocontrol systems primarily because they want to have system responsibility and a better understanding of the necessity for high control loop bandwidths. The cost of greater bandwidth is not only in the increased complexity of associated control equipment, but in the size and strength of the machines and actuators which must withstand the greater velocities and accelerations associated with larger bandwidths.

The panel moderator, Mr. Chestnut, added to Dr. Bower's observations by noting that other limitations in machine tool performance are in the sensors and on-line analyzers, and that a real obstacle in nonlinear control problems is the difficulty in measuring the controlled variable except by methods which involve considerable time. He indicated that some controlled variables in chemical processes are actually measured after the process has been controlled.

The question was asked whether anyone had information about on-line analyzers.

Dr. Taylor commented that all the examples of online analyzers seemed to be in the chemical industry, and he gave as an example the VPC or Vapor Phase Chromatograph. He also enlarged on the subject by discussing the importance of resolution and accuracy of the servos. With end-point analyzers having good resolution but not absolute accuracy, an accurate control system could be built, he believed, by introducing small variations in a digital controller, and by using the resulting variations in the outputs (measured by the end-point analyzers) to adjust the set points.

A question from the floor induced Dr. Taylor to clarify his last statement somewhat by remarking that the feedback device in the control system would be a computer. However, he stated that if the instrument manufacturers would make their sensors more accurate it would not be necessary to incorporate such a costly device.

A member of the audience noted that on-line computing is being done in large wind tunnel research to eliminate expensive stops which would otherwise be needed to evaluate the test results.

After this statement, Mr. Chestnut closed the open panel discussion by saying, "Well, I am sure this is a subject that we could not exhaust were we to continue for another 12 hours, and I am sure we would be exhausted long before then. So by way of summary now that the facts are in, let each of our members of the panel comment on the obstacles as they originally started out, and as they now feel about them in light of the over-all discussion which has taken place."

Closing Remarks

THE panel proceedings and the problems of nonlinear controls were aptly summarized by the panel members. Note that some of their remarks were influenced by the previous discussions, and that their conclusions are not necessarily in complete accord.

Prof. Smith

Well, I'd like to comment not on an obstacle but on a way of overcoming an obstacle. I think that one of the greatest contributions that we can make in the nonlinear field is using the old computer back here (our brains) and inventing new modes of control. You can't do this by just solving for that nonlinearity which, placed in series with a gain and some other nonlinearity in a single loop, would produce the type of control you want. Because, a single loop is incapable of those thousands of other types or modes of control that different topological patterns will give you. You either have to follow the guide of your mathematics to find these new topological patterns, or you can use a little ingenuity and drive your control system as a man throws a football or drives a car around a rapid curve. He swings with the car and hits it twice instead of once in order to have two vectors to cancel out the oscillation of the car, and this sort of ingenious concept of a new mode of control is what you need to have in order to have better control systems.

Dr. Bower

Since I am beyond correction now by Hal Chestnut, I will return to the mathematical problem. I think the only real obstacles are going to remain in our own ignorance. I like to recall the comment made by Einstein when he was a young man and was considering entering

the field of mathematics as his life career, and decided to give it up right then and there because it would take him too long to learn it. I think to some extent this is true of the engineer who realizes now that he needs a terrific amount of mathematics to handle the automatic control problem adequately and I would like to vote for this idea of simplifying the access of the engineering student to mathematics. I'm not suggesting another book which summarizes higher mathematics for the engineer, but I am suggesting some real attention to, you might say, the means of opening to the engineer just a brief glimpse of new fields of mathematics as they develop. He will have to work by himself in these fields once he discovers their applicability, but, he needs a means of finding out what is available.

Mr. Holzmann

Looking at the problem of nonlinear control from the point of view of the engineer concerned with chemical refining or nuclear process plants, it would seem that progress has been limited not so much by the mathematical difficulties as by the economics of the situation.

Emphasis today is on analysis and synthesis. When a problem leads us beyond available theory, we must do what engineers have always done—use imagination, judgment, and patience to arrive at a trial-and-error solution. The old-timers in the process control field devised nonlinear control schemes which have operated successfully for many years. Their approach was essentially practical. By the way, their accomplishments have rarely found recognition in the technical press. It was taken for granted that they would find a solution, no matter how tough the problem at hand.

Analysis can help us to separate good design ideas from poor ones, and save development time and money. Sometimes mathematical analogy becomes a creative tool, suggesting new avenues of attack on a problem. But analysis is not the only pathway to progress. In fact, the analysis of any process control problem is liable to arrive at a point where further refinement pays diminishing returns. Beyond this point, one may find it more economical to shift the emphasis from desk to shop and laboratory.

We have not had much discussion on the subject of economics, possibly because the majority of the audience here work in areas where economics does not seem to be a problem. In the process industries it appears to be the prime obstacle to progress in nonlinear control. Progress will not lag once management in these industries becomes convinced of the wisdom of channeling more funds into research and development of advanced control systems.

Mr. West

I feel that one of the obstacles to progress in nonlinear control is the lack of an adequate mathematical theory. While it is true that computers can be employed to predict the performance of a particular control system, a better mathematical theory is required before the computer can be utilized effectively to synthesize a desired response.

I don't feel that economic considerations are a bar to the use of computers in control applications. There are applications in which a general-purpose computer can be used to periodically set the operating points of existing process controllers. The control established by the computer will be enough better than the manual control now employed to justify acquisition of the system. By carefully selecting the first such applications, it will be possible to show an initial savings on the system with a simple mathematical model of the process. As experience is gained, the computer program can be revised to exercise better control, while employing the same process controllers or to exploit improved controllers as they become available. The system is flexible and can be easily modified to assume additional control functions or to tighten the control through improved computational procedures. It probably will be possible to find applications in which the computer does an adequate job of controlling the process and, in addition, records information about the process which can be used to improve the method of control or to bring additional variables under the control of the computer. The difficulty is in selecting the initial applications so that there will be an economic advantage with the first control program which may be just slightly better than manual control. Experience with the process will almost certainly allow tighter control and a better economic justification.

Dr. Taylor

Well, I'm for cats and the cat experiment. This may sound facetious but it's really meant to indicate that I'm for marrying many fields together. Medical electronics, the ties between—the whole idea here is that I believe there is a whole host of helpful approaches in various fields and it's got to be a matter of coming together sideways and going up together. Now, tied in to this is economic justification, and I'm afraid I have to take some exception with George West. If you have ever gone out to try to sell management on digital controlling in the process industry or something like that, I know you have quickly run into management-economic-justification-type barriers and these will only come down in time. But in the meanwhile, progress, I think, is going to best go ahead where companies invest their own time and money and sweat. I've been trying to emphasize that in the military game funds are a little easier to come by than when you try to hit the industrial trail with new approaches and new ideas in the control field.

Mr. Chestnut

Our program here has emphasized the obstacles, perhaps, and ways to better understand our programs associated with progress in nonlinear controls.

Actually, I am quite optimistic in terms of the progress that will be made and is being made in the nonlinear control field, I think the attention that IRE is giving here, in focusing the minds of a large group at this Symposium, and in the next few days at WESCON, is a very helpful step in emphasizing our problem and giving us something further to work on.

I think that a large number of tools are already available. I'm not saying that they are the ultimate tools, but there are some very powerful ones and it would appear as though we can make considerable progress with what we have. And I am sure that with improvement in tools available, both mathematical as well as others, we will be able to move forward rapidly.

The last thing I am optimistic about is that when the facts of the dollars are known, I think that economic forces will force the use of nonlinear control. I think, as Mr. Holzmann has pointed out, the operations research study of a number of cases will show that potential controls users are losing money by not spending the money.

However, I think it is going to be up to engineers to help provide the data to the operations research people.

Well, with this as a concluding remark, I would like to say that it has certainly been a privilege on my part to be a member of this panel, and I want to thank the panel members and those in the audience for their fine participation.

COMMITTEE FOR THE 1957 PGAC SYMPOSIUM ON NONLINEAR CONTROL

Chairman
E. M. Grabbe
The Ramo-Wooldridge Corp.

Vice-Chairman
John Lozier
Bell Telephone Laboratories

Technical Program
T. M. Stout
The Ramo-Wooldridge Corp.

WESCON Coordination
A. M. Hopkin
University of California

AIEE Coordination
J. G. Truxal
Brooklyn Polytechnic Institute

ASME Coordination
W. E. Vannah
McGraw-Hill Publishing Co.

Arrangements
H. S. Robinson
Westinghouse Electric Co.
Frank Muller
University of California
Peter Rogers
University of California
P. E. Merritt
Stanford Research Institute

Publicity
J. M. Jones
Hughes Aircraft Co.

Luncheon
Elliot Lewis
The Ramo-Wooldridge Corp.

SPONSORED BY THE
PROFESSIONAL GROUP ON AUTOMATIC CONTROL
OF THE
INSTITUTE OF RADIO ENGINEERS
with participation by
Feedback Control Systems Committee
of the
American Institute of Electrical Engineers
and
Instruments and Regulators Division
of the
American Society of Mechanical Engineers

Book Review

"Analytical Design of Linear Feedback Controls," G. C. Newton, Jr., L. A. Gould, and J. F. Kaiser. New York: John Wiley and Sons, Inc., 1957, 419 pp., $12.00. *Reviewed by J. E. Gibson*†

This book should be of considerable interest to the controls engineer. It is an introduction to system design by the method of analytical optimization of system parameters to a given criterion. It is definitely not just another in the series of books on basic servomechanism theory. Although James, Nichols, and Phillips and Laning and Battin (essentially reference books) have preceded it in the general area, this is the first book on the text level to treat this subject. There are approximately 280 pages of text and 100 pages more of appendixes. The 280 pages are an excellent but expensive twelve dollars' worth.

The text begins the study of analytic design with linear systems, typically subjected to step inputs. It then advances to the system with stochastic inputs and then to certain types of nonlinear systems. The optimization criterion is rms error or integral square error. By the use of Parseval's theorem, the authors are able to determine the integral square value of a time function in terms of its Laplace transform, thus sidestepping the problem of direct and inverse transformation while retaining the convenience of the transform mathematics. The text is well organized and presents an integrated treatment of the method. No previous familiarity with random variables is assumed and the necessary background on stochastic processes is presented clearly and simply. The authors wisely use very simple examples in the body of the text so that the philosophy is not obscured by calculation details. The problems for each chapter are somewhat more realistic; in Chapter 9 a detailed example design of a practical system is discussed. The text is certain to become a standard in this area.

Since the method proposed by this text is quite different from the usual methods, a comparison of the two approaches is in order. In this reviewer's judgment, the "analytic" method proposed by the authors is not worth the trouble involved when applied to linear systems with continuous inputs. The authors were perfectly justified in starting their exposition with continuous input systems to acquaint the reader with the method. However, they attempt to justify their method on its merits rather than on pedagogic grounds. On pages 29–31, the authors point out the disadvantages of the usual "cut-and-try" method of design as compared to their "analytic" method. There are several points they overlook. First, analytic approaches to the conventional methods do exist, but engineers have found that cut and try is somewhat faster. In other words, cut and try is a matter of choice, not of necessity. Second, there is no reason to finally reach a design that has excessive bandwidth by conventional methods, as the authors imply. And finally, there is no reason why the choice of compensation need be purely a matter of experience, as the authors state.

Next we come to an "advantage" of the analytic method—the use of the rms error criterion or integral error criterion. The usual argument for the use of these measures of system performance is one of conveniene or mathematical simplicity and not that they are particularly good figures of merit. The work of Graham and Lathrop[1] on the ITAE specification[2] gives an excellent discussion of measures of system goodness and rms error ranks well down on their list. On page 84, Newton, Gould, and Kaiser point out that even with these criteria, the "analytic" method becomes mathematically unwieldy in many practical situations.

On page 41, the authors state: "Since the minimization of the integral square error can be done analytically for a rather general class of signals, its use is justified very frequently as a substitute for performance indices which are more closely associated with the design objectives, but which lead to insolvable mathematical problems." Thus, apparently we are told that if you can't solve the problem you want to solve, solve the problem that you can solve. (Not that this is an unusual engineering approach!) Then finally on page 45, "If the equations for the parameters obtained by setting the partial derivatives of the integral-square error with respect to the parameters equal to zero are too complex to be solved by analytical methods, we may resort to *trial and error* or *successive approximation* techniques . . . ." (The italics are this reviewer's.) Have we not, then, come the full circle?

The authors' enthusiasm for their method is understandable, but their judgment of the conventional "try-and-error" method seems unduly harsh and their judgment of the practicability of their "analytic method" appears unduly optimistic.

The real worth and importance of this first-class work is to introduce the reader to the more elaborate mathematical techniques that are necessary when the designer is faced with stochastic inputs; the greater portion o the work is concerned with this subject After considering linear systems subjected to stochastic inputs, the authors extend their discussion to systems subjected to saturation. They approach the problem of compensating linear systems by solving for the weighting function of a compensation element which yields minimum rms error or integral squared error for a particular input. The input may be a stochastic signal or a standard transient such as a ramp or step. The result of the analysis yields an integral equation of the Wiener-Hopf type. While the mathematician calls this "reducing the problem to quadrature" and considers it solved, the engineer must continue on to instrument a simple, passive if possible, physically realizable compensation network from the equation. This generally is not simple. But at the present time, it is the only available method for optimizing a system when the input statistics are to be taken into consideration.

This book will serve as an excellent introduction to the design of control systems from the modern communications theory point of view. Due to the inadequacy of the rms criterion, however, and the difficulty of making assumptions about the noise contamination that are at once justified and mathematically tractable, the design problems are extremely difficult. While this text is an excellent start, there remains much more to be done before the application of modern theories of communication will result in a significant improvement in the performance of any given physical control system. Even the classic example of the radar gun control, with its contaminated return signal, is not improved appreciably by these techniques, as the reader may see from page 328 of James, Nichols, and Phillips.

The authors of the present work are to be congratulated on their efforts to refine the design of automatic control devices and make it more scientific. They have been successful in reducing the design of systems with stochastic inputs from what seemed to some an esoteric rite of higher mathematics to an engineering design method. The authors represent a considerable segment of opinion in the automatic control field in their dissatisfaction with the present state of the design art.

It remains to be seen whether the trend in new design methods will be toward the more elaborate, analytic methods represented by this text, or on the other hand, a return to the most primitive methods of examining the roots of the characteristic equation and manipulating these roots to obtain a desired response—all done automatically by digital computers, of course.

† Assoc. Prof., School of Elec. Eng., Purdue University, Lafayette, Ind.

[1] D. Graham and R. C. Lathrop, "The synthesis of optimum transient response: criteria and standard forms," *Trans. AIEE (Applications and Indus.)*, vol. 72, pt. 2, pp. 273–288; November, 1953.

[2] ITAE represents Integral of Time-multiplied Absolute-value of Error.

PGAC News

1958–1959 PGAC Administrative Committee

The following officers and members were elected at the annual committee meeting in March, 1958.

Chairman

J. E. Ward, Massachusetts Institute of Technology, Cambridge, Mass.

Vice-Chairman

J. M. Salzer, The Magnavox Co., Los Angeles, Calif.

Secretary-Treasurer

G. A. Biernson, Sylvania Electric Products, Inc., Waltham, Mass.

Members

G. S. Axelby, Westinghouse Electric Corp., Baltimore, Md.
Victor Azgapetian, Servomechanisms, Inc., Westbury, N. Y.
N. H. Choksy, Johns Hopkins University, Baltimore, Md.
E. M. Grabbe, The Ramo-Wooldridge Corp., Los Angeles, Calif.
Harold Levenstein, W. L. Maxson Corp., New York, N. Y.
D. P. Lindorff, University of Connecticut, Storrs, Conn.
J. C. Lozier, Bell Telephone Laboratories, Inc., Whippany, N. J.
T. F. Mahoney, Raytheon Manufacturing Co., Wayland, Mass.
J. H. Miller, Goodyear Aircraft Corp., Akron, Ohio
O. H. Schuck, Minneapolis-Honeywell Regulator Co., Minneapolis, Minn.
O. J. M. Smith, University of California, Berkeley, Calif.
T. M. Stout, The Ramo-Wooldridge Corp., Los Angeles, Calif.
A. R. Teasdale, Jr., Temco Aircraft Corp., Dallas, Texas
R. B. Wilcox, Sylvania Electric Products, Inc., Waltham, Mass.
Felix Zweig, Yale University, New Haven, Conn.

1959 National Automatic Control Conference

The PGAC will sponsor a National Automatic Control Conference on November 4–6, 1959, in Dallas, Texas. The convention will be held in the Sheraton Dallas Hotel, scheduled for completion in March, 1959.

The Feedback Control Systems Committee of the AIEE and the Instruments and Regulators Division of the ASME will participate in the activities.

Although the deadline for papers will not take place until the summer of 1959, abstracts and summaries of proposed papers should be sent as soon as possible to the Editor of PGAC Transactions:

George S. Axelby
Air Arm Division
Westinghouse Electric Corporation
Box 746, Baltimore 3, Md.

More details on the conference will be announced in the coming months.

Dr. William G. Tuller Memorial Award

Awards of $250.00 in memory of Dr. William G. Tuller will be presented for the best papers written by senior or graduate students on the subject of component parts. The topic may relate to operational theory, materials, construction, design, testing, or application of any electronic component.

Papers must be submitted by December 31, 1958. For additional information consult the Office of the Dean or the IRE Faculty Advisor.

INSTITUTIONAL LISTINGS

The IRE Professional Group on Automatic Control is grateful for the assistance given by the firms listed below, and invites application for Institutional Listings from other firms interested in the field of Automatic Control.

The charge for an Institutional Listing is $75.00 per issue or $125.00 for two consecutive issues. Applications for Institutional Listings and checks (made out to the Institute of Radio Engineers, Inc.) should be sent to Mr. L. G. Cumming, Technical Secretary, Institute of Radio Engineers, Inc., 1 East 79th Street, New York 21, N. Y.